ICACT

INTERNATIONAL CONFERENCE ON ADVANCED COMPUTING TECHNOLOGIES | 2024

NSBM GREEN UNIVERSITY TOWN

PROCEEDINGS

EMPOWERING DIGITAL TRANSFORMATION THROUGH ADVANCED COMPUTING INNOVATIONS

OCTOBER 24, 2024

RESEARCH CONFERENCE 2024 - ORGANIZED BY FACULTY OF COMPUTING

THE 01ST INTERNATIONAL CONFERENCE ON ADVANCED COMPUTING TECHNOLOGIES

"EMPOWERING DIGITAL TRANSFORMATION THROUGH ADVANCED COMPUTING INNOVATIONS"

FOCUS AREAS:

- Data Science & Artificial Intelligence
- Computer Networks, Cyber Security, and Computer Systems
- Computer Vision, Image, and Signal Processing
- Information Systems, HCI, and Enterprise Systems
- Smart Living, Digital Technologies and IoT
- Social Media and Web Technologies

Organized by

Faculty of Computing,
NSBM Green University
Pitipana, Homagama,

Disclaimer

The views expressed in the papers published in these proceedings are those of the authors. The statement and opinions stated either in the papers published herein or at the conference and discussions do not necessarily represent the views of the NSBM Green University of Sri Lanka. All papers are formatted using the seventh edition APA referencing style. If an author wants to present a work that has already been published or accepted by the conference, they need to write to the chair of the conference and inform that they are publishing the work elsewhere. However, the authors have the authority to present only a part of the study and not the content in its entirety. There might be some restrictions on how the authors can reuse the material that has been published/accepted.

ISSN 3051-5106
ISBN 978-955-23-0019-3

Conference Agenda

10:00 am	Registrations
11:00 am	Lighting of the Traditional Oil Lamp & Welcome
11:15 am	Welcome Speech Prof. Chaminda Wijesinghe, Dean – Faculty of Computing
11:20 am	Conference Proceedings Presenting
11:30 am	Keynote Speech 1, Prof. Nathan Clarke, Professor of Cyber Security and Digital Forensics
11:50 am	Entertainment Act – Conducted by NSBM Music Club
12:00 pm	Keynote Speech 2, Dr. Malith Jayasinghe, Vice President Research & AI, WSO2.
12:20 pm	Concluding Remarks, Dr. Mohamed Shafraz, Conference Chair
12:30 pm	Lunch
01:30 pm	Commencement of the Conference Parallel Sessions

PREFACE

We are pleased to present the proceedings of ICACT 2024 and extend a warm welcome to all authors and delegates attending this prestigious event. We are confident that the insights within will be both enlightening and transformative. This year marks a special milestone as the Faculty of Computing at NSBM Green University proudly hosts the International Conference on Advanced Computing Technologies (ICACT 2024) for the first time, pushing the frontiers of academic exploration. NSBM continues to lead in higher education by fostering innovation and promoting scholarly exchange.

ICACT 2024, under the theme "Empowering Digital Transformation through Advanced Computing Innovations" boldly ventures into the rapidly evolving technological landscape. In an age where digital advancements are reshaping industries and societies, this conference seeks to ignite discussions on the role of technology in driving meaningful and positive change. Emerging technologies have become essential for organizations and nations, presenting new opportunities and challenges at the intersection of business and technology.

This conference serves as an important platform to unite local and international researchers, academics, and industry leaders in exploring the profound effects of digital transformation on society. The chosen theme underscores the urgent need to harness digital innovation for social progress. Through strategic implementation, businesses, governments, and communities can work together to accelerate positive societal change via emerging technologies.

As we embark on this intellectual journey, we extend our heartfelt gratitude to Prof. E. A. Weerasinghe, Vice-Chancellor of NSBM Green University, and Prof. Chaminda Rathnayaka, Deputy Vice-Chancellor, for their unwavering support and visionary leadership. Special thanks go to Prof. J. Baratha Dodankotuwa, Head of Academic Development and Quality Assurance, for his valuable contributions, and to Prof. Chaminda Wijesinghe, Dean of the Faculty of Computing, for his expert guidance and knowledge.

We sincerely appreciate all the local and international presenters and participants whose contributions enhance the quality of this conference. ICACT 2024’s success is a result of the hard work and dedication of our reviewers and the diligent efforts of the Conference Organizing Committee.

In conclusion, we expect ICACT 2024 to spark rich academic discussions and generate innovative solutions that contribute to a digitally transformed and progressive future. We wish all attendees a rewarding and memorable experience.

The Conference Organizing Committee
ICACT 2024

MESSAGE FROM VICE-CHANCELLOR

Prof. E.A. Weerasinghe
Vice Chancellor

It is with immense pleasure and pride that I warmly welcome you to the International Conference on Advanced Computing Technologies (ICACT) 2024, hosted by the Faculty of Computing at NSBM Green University. As we launch this inaugural research conference, we find ourselves at the intersection of academia, innovation, and the transformative potential of digital technologies.

Our theme for ICACT 2024, "Empowering Digital Transformation through Advanced Computing Innovations" highlights the ever-evolving landscape of research and innovation. In today's fast-paced global environment, the seamless integration of emerging technologies with business processes is not just relevant—it is essential for driving societal progress. ICACT 2024 builds on the success of our previous initiatives, serving as a beacon for scholars, researchers, and industry experts to come together and examine the profound impacts of digital transformation. We stand on the brink of a new era, where the fusion of advanced computing technologies with business operations holds the potential to streamline processes and shape a future that is both sustainable and inclusive.

Universities play a crucial role in driving research and fostering innovation, contributing to economic and societal development. ICACT 2024 exemplifies this commitment by providing a platform for multidisciplinary researchers to share their work and engage in productive discussions on how digital technologies can empower businesses to operate more efficiently and contribute to societal growth. I extend my heartfelt congratulations to the organizing committee for their dedication to crafting a conference that promises to be a wellspring of knowledge, ideas, and innovative solutions. The high-quality papers and expert insights presented at this event are a reflection of our academic community's commitment to advancing the boundaries of knowledge.

A special thanks goes out to all the presenters and delegates whose contributions enrich our conference with insightful presentations and meaningful dialogues. Your participation plays a key role in strengthening our university's research culture and, more importantly, in supporting broader national development. I eagerly anticipate the stimulating discussions that will take place at ICACT 2024. May this conference serve as a source of inspiration, collaboration, and transformative ideas, propelling us toward a future where digital transformation not only enhances business processes but also accelerates positive societal change.

Thank you!

MESSAGE FROM DEPUTY VICE-CHANCELLOR

Prof. Chaminda Rathnayake

Deputy Vice Chancellor

It is with great pleasure that I extend my warmest greetings to the inaugural International Conference on Advanced Computing Technologies (ICACT 2024), hosted by the Faculty of Computing at NSBM Green University. As the Deputy Vice-Chancellor, I am honored to welcome you to this esteemed event centered around the theme "Empowering Digital Transformation through Advanced Computing Innovations".

In a time of remarkable technological progress, the influence of digital transformation on various aspects of society is both profound and wide-reaching. ICACT 2024 provides an essential platform for sharing ideas, insights, and research that deepen our collective understanding of the transformative role of digital innovation. At NSBM Green University, we are committed to nurturing an environment that promotes academic excellence, innovation, and collaboration. This conference perfectly aligns with our mission to encourage cutting-edge research and stimulate discussions that contribute to the advancement of business practices and societal well-being.

As we engage in discussions focused on empowering digital transformation through advanced computing, I encourage each of you to actively participate, share your knowledge, and explore opportunities for collaboration. The diverse viewpoints represented here will undoubtedly enhance our understanding of the challenges and opportunities that emerging technologies bring, especially in how they streamline and optimize business processes.

I would like to extend my sincere appreciation to the organizing committee, presenters, reviewers, and all participants for their dedication and efforts in bringing ICACT 2024 to life. Your commitment to advancing knowledge and driving positive change is truly commendable, and I am confident that this conference will foster meaningful developments in the field of business innovation through technology. May your experiences at ICACT 2024 be both intellectually stimulating and personally rewarding. I wish you a highly productive and inspiring conference.

Thank you!

MESSAGE FROM THE HEAD OF ACADEMIC DEVELOPMENT AND QUALITY ASSURANCE

Prof. J Baratha Dodankotuwa

Head Of Academic Development and Quality Assurance

It is my distinct pleasure to extend a warm welcome to all participants, researchers, and scholars who have gathered for the inaugural International Conference on Advanced Computing Technologies (ICACT 2024), proudly hosted by the Faculty of Computing at NSBM Green University. As the Head of Academic Development and Quality Assurance, I am thrilled to witness the assembly of brilliant minds and groundbreaking ideas converging around the central theme of "Empowering Digital Transformation through Advanced Computing Innovations".

The rapidly evolving digital landscape is not only revolutionizing academia and industry but also fundamentally transforming the way we live and work. ICACT 2024 provides an exceptional platform for us to delve into the many ways in which emerging technologies—such as artificial intelligence, big data analytics, cloud computing, and blockchain—are driving this digital revolution. These technologies hold the potential to reshape business processes, making them more efficient, agile, and adaptable to the challenges of the modern world. The conference encourages us to explore how leveraging such technologies can streamline operations, optimize decision-making, and create new opportunities for growth and innovation.

At NSBM Green University, we are dedicated to upholding the highest standards of academic excellence. Conferences like ICACT play a critical role in fostering a culture of inquiry, collaboration, and continuous improvement. As we embark on this intellectual journey, I urge you to actively engage in the sessions, present your research, and participate in insightful discussions with your peers. The theme of this conference, focusing on digital transformation, emphasizes the need for businesses and societies to adapt swiftly and strategically to the digital era. By embracing advanced computing technologies, businesses can smooth out complex processes, reduce inefficiencies, and accelerate their path to success.

I would also like to take this opportunity to express my heartfelt gratitude to the organizing committee, and all contributors who have worked tirelessly to bring ICACT 2024 to life. Your unwavering dedication to advancing knowledge and fostering innovation is pivotal to the success of this conference. It is our hope that the knowledge shared, and the collaborations formed here will deepen our understanding of the transformative power of digital technologies, while also highlighting their ability to drive business innovation and societal progress. May your experience at ICACT 2024 be intellectually enriching and may the connections you make during this conference lead to fruitful partnerships that further propel the boundaries of technology-driven business and societal advancements.

Thank you!

MESSAGE FROM THE DEAN

Prof. Chaminda Wijesinghe
Dean, Faculty of Computing

Dear Students, Faculty Members, and Esteemed Guests,

It is with immense pride and joy that I extend my heartfelt congratulations to all participants and organizers of the first-time conference on International Conference on Advanced Computing Technologies (ICACT 2024) of the Faculty of Computing at NSBM Green University. Today marks a significant milestone in our academic journey, showcasing the innovative spirit, dedication, and intellectual prowess of our students.

The establishment of this first-ever conference is a testament to our commitment to fostering a robust research culture within the Faculty of Computing. It provides a unique platform for our students to present their pioneering research, engage in scholarly discussions, and collaborate with peers and experts in the field. Your hard work, curiosity, and passion for knowledge have brought us to this remarkable moment, and I am confident that the insights and discoveries presented today will pave the way for future advancements in computing and technology. I would like to express my gratitude to the faculty members, reviewers, and staff whose guidance and support have been instrumental in the success of our students' research endeavors. Your unwavering dedication to nurturing the next generation of innovators is truly commendable. To our students, I commend you for your perseverance and creativity. Your contributions are not only significant to your personal growth but also to the broader field of computing. As you present your research, remember that this symposium is just the beginning of your journey as a researcher and thought leader.

Congratulations once again to all participants for your outstanding achievements. I wish you a productive and inspiring symposium. Let us celebrate this momentous occasion and look forward to many more successful research symposia in the years to come.

Thank you!

MESSAGE FROM THE CONFERENCE CHAIR

Dr. Mohamed Shafraz
Conference Chair, ICACT 2024

I am deeply humbled and honored to share this message for the first-ever International Conference on Advanced Computing Technologies (ICACT) 2024, which will take place on 24th October 2024 at NSBM Green University. As South Asia's first green university, NSBM Green University continuously strives to make meaningful contributions to the nation and the global community through research and development, embodying our essential role in advancing academic endeavors.

The theme of ICACT 2024, "Empowering Digital Transformation through Advanced Computing Innovations," highlights the vital role of digital innovation in today's world. The conference aims to explore how technologies like AI, machine learning, IoT, and cloud computing are reshaping sectors such as healthcare, education, business, and governance. These emerging technologies are no longer confined to IT but drive efficiencies, optimize decision-making, and solve complex societal challenges across industries. This conference serves as a platform to explore the transformative potential of these innovations and how they can accelerate not only business success but also social and economic progress. In making ICACT 2024 a reality, countless individuals have contributed to every phase of the process.

This achievement would not have been possible without the visionary leadership of our Vice-Chancellor, Prof. E.A. Weerasinghe, whose guidance and encouragement have been a constant source of strength. I would also like to extend my heartfelt gratitude to Prof. Chaminda Rathnayake, Deputy Vice-Chancellor, and Prof. Baratha Dodankotuwa, Head of Academic Development and Quality Assurance, and Prof. Chaminda Wijesinghe, Dean of the Faculty of Computing for their immense support in enhancing the research culture at NSBM.

On behalf of the Organizing Committee, I would like to express a special note of gratitude to all local and international speakers, authors, reviewers, researchers, and presenters for their invaluable contributions to ICACT 2024. Your time, efforts, and insights are what make this conference a success. As the Conference Chair, I encourage all participants to actively engage in discussions, share their valuable expertise, and seek collaborative opportunities. The collective knowledge and expertise gathered here have the potential to drive significant change. Let us seize this moment to shape the digital transformation narrative, ensuring it aligns with our shared values, enhances business processes, and propels us toward a future where technology is a powerful force for positive societal change. Thank you.

ORGANIZING COMMITTEE

ADVISORY BOARD:
Prof. E. A. Weerasinghe - Vice Chancellor
Prof. Chaminda Rathnayaka - Deputy Vice-Chancellor
Prof. J. Baratha Dodankotuwa - Head, Academic Development and Quality Assurance
Prof. Chaminda Wijesinghe - Dean, Faculty of Computing
Prof. Noel Fernando - Professor, Faculty of Computing

CONFERENCE COMMITTEE:
Dr. Mohamed Shafraz - Conference Chair
Dr. Pabudi Abeyrathne - Conference Co-Chair
Mr. Chamara Dissanayake - Conference Co-Chair
Ms. Lakni Peiris - Conference Secretary

REVIEW BOARD:
Prof S.Vasanthapriyan
Prof. Nalinda Somasiri
Prof. Noel Fernando
Prof. Chaminda Wijesinghe
Prof Chaminda Wijesinghe
Dr. Dilrukshi Gamage
Dr. Kushani De Silva
Dr. Gihan Chathuranga
Dr. Hiruni Rupasinghe
Dr. Nirosha Wedasinghe
Dr. Shan Jayasinghe
Dr. Manoja Weerasekara
Dr. Pabudi Abeyrathne
Dr. Rasika Ranaweera
Dr. Mohamed Shafraz
Mrs. Udeshika Indumini
Ms. Dilhara Batanarachchi
Mr. Shashika Suren
Ms. Nimasha Arambepola
Mr. Indrajith Jayasuriya
Mr. Liyanage Perera
Ms. Chalani Oruthotaarachchi
Mr. Chamara Dissanayake
Mr. Chamara Attanayake
Ms. Pavithra Subashini
Mr. Isuru Sri Bandara
Ms. Hirushi Dilpriya
Mr. Gayan Perera

COMMITTEE MEMBERS:
Mr. Naji Saravanabavan
Mr. Chamindra Attanayake
Ms. Pavithra Kankanamge
Ms. Dulanjali Wijesekara
Ms. Natasha Chamba
Mr. Madusanka Mithrananda
Ms. Yasanthika Mathotaarachchi
Ms. Isuri Caldera
Mr. Gayan Perera
Ms. Nethmi Weearsinghe
Ms. Hirushi Dilpriya
Mr. Isuru Sri Bandara
Ms. Kavishka Rajapakshe
Ms. Sarangi Attanayake
Ms. Chathurma Wijesinghe
Ms. Chathurika Ranaweera
Ms. Gayathmi Kariyapperuma
Mr. Vimukthi Nilupul
Mr. Poorna Hettiararchchi

EDITORIAL COMMITTEE:
Mr. Anton Jayakody
Ms. Thisarani Wickaramasinghe
Mr. Diluka Wijesinghe
Ms. Sanuli Weerasinghe
Ms. Dharani Rajasinghe
Mr. Supun Gajendrasinghe

KEYNOTE SPEECH I

Prof. Nathan Clarke

Abstract: Cybersecurity continues to be a significant challenge with reports stating cybercrime will cost the global community $10.5 trillion by 2025. Cybercrime is not a new problem, but the scale is escalating to new levels, and the impact cuts across all of society. Artificial intelligence (AI) is seen by many as a potential step change in capability for cyber security - yet the prospect of truly intelligent systems also opens up equal opportunities for the hacker community. This presentation will draw upon Prof Clarke's 20+ years of developing AI and machine learning systems to aid in a variety of cyber security solutions - providing examples of where and how it can be used to provide more effective cyber security. The talk will also address the weaknesses of current approaches and the need for a new generation of explainable and trustworthy AI to ensure we understand how these systems operate to protect us. Finally, Prof Clarke will draw upon some examples of how technology can be misused and the potential dangers that exist for us all now and in the future.

KEYNOTE SPEECH II

Dr. Malith Jayasinghe

Abstract: Software engineering is a fundamental discipline within computer science, focusing on the systematic design, development, testing, and maintenance of software systems. Recent advancements in artificial intelligence (AI) have significantly transformed this field, leading to remarkable improvements in the software development process. AI enhances efficiency by automating tasks such as code generation, debugging, testing, and system maintenance through intelligent tools. However, it also brings forth several challenges, including ensuring accuracy, maintaining acceptable speed, protecting data privacy, addressing security vulnerabilities, and managing the inherent complexities of AI-driven systems. In this talk, I will explore AI's substantial impact on software engineering, the challenges associated with developing AI-enabled tools that foster these innovations, and strategies to overcome those challenges.

Table of Contents

Web Application for Predicting and Summarizing Medical Test Report Results Using Machine Learning Algorithm

H N Semage
Faculty of Computing
NSBM Green University
Mahenwaththa, Pitipana, Homagama, Sri Lanka
himasha.semage@gmail.com

***Abstract*— Healthcare domain is an important domain which exists in society. In Healthcare domain health literacy is a major factor. Health literacy amon g patients on medical test report results has impacted in a major way to the health of the patients. This factor must be addressed since health literacy among patients on medical test report results impact the decision-making process and the understanding of the patient's help. In this work, the aim is to address the problem of understanding the medical test report results by patient on their own. The proposed project uses Supervised Learning algorithm. The primary object is to provide with an accurate prediction for the medical test report results and provide details based on the prediction given like, what the prediction means, symptoms, risks, precautions, and provide URL links to get more idea about the reports by developing a web application. Before selecting a suitable algorithm, few selected algorithms were evaluated based on the accuracy and for the project with 100% accuracy for training data and 97.50% accuracy for testing data, Random Forest is used for predicting the prediction of the medical test report results. The algorithm is tested by applying datasets.**

Keywords—Healthcare domain, Health literacy, Web application, Supervised Learning Algorithm, Random Forest algorithm, Accuracy.

I. INTRODUCTION

Healthcare industry has a major impact on society, since this is the area of knowledge and practice that deals with maintains, improvements and enhancement of individuals' and populations' physical health, mental health, and social well-being. This is very sensitive domain which the results impact on one's life. This vast area of health care covers prevention, diagnosis, treatment and management of illnesses, injuries, and disabilities and also promotes healthy lives and activities. There are different resources such as patients, doctors, nurses, laboratory staff and equipment, that play different roles which work together in order to achieve desired and required outcomes.

In healthcare domain, health literacy is a main factor that should be considered. Health literacy is the ability of an individual to find, understand and use information and services to inform health-related decisions and actions for themselves and others [1]. Lack of health literacy is mostly seen in patient. In an article it has been explained how healthy literacy has been affected to patients with heart failures – "Among patients with heart failure in an integrated managed care organization, low health literacy was significantly associated with higher all-cause mortality" [2]. One of the most common factors when it comes to lack of health literacy is that patient's struggle to understand medical test report results. Because of lack of health literacy patients tend to listen to medical professionals without any proper idea what will be the result of the decisions that they are taking. An article talks about the ap between physicians and patients [3]. This might lead to negative impact and risks on an individual's health [4].

Different machine learning algorithms were applied to find the best algorithm for the prediction of medical test report results. The main objective of this paper is to suggest a web application that provides predictions and other information related to the prediction to resolve lack of health literacy in patients regarding medical test report results, while finding the best fit algorithm among different machine learning algorithms. By the prediction the patients will understand their level of health condition and other information related to the prediction like what the prediction means, symptoms, risks, precautions and provide URL links to search more information about the medical reports.

II. LITERATURE REVIEW

This research includes the study of different techniques used for prediction of medical test report results. There is no existing system found for the exact system that is suggested but there are existing systems that use same concept of the suggested system. For this purpose, different papers have been studied to get information about the algorithm and workflow that is used in similar existing systems.

A. Red Wine Quality Prediction using Machine Learning Techniques

This discusses a model given for predicting red wine quality wine machine learning techniques. Three algorithms were used to analyze and get the best of the three algorithms based on n the results of the training set after various measures have been calculated, and the results are compared between the training and testing sets. Based on the results obtained they have chosen Support Vector Machine as the machine learning algorithm to the model [5]. The workflow of the model focusses mainly on evaluating the three algorithms based on the measures by using training and testing data and finding the best fit algorithm to the model. As the conclusion Support Vector Machine algorithm is chosen as the best fit algorithm to the model.

B. Predicting the Survival Rate of Titanic Disaster using Machine Learning Approaches

Predicting the survival rate of Titanic disaster using machine learning approaches discusses about the model suggested to predict the survival rate of titanic disaster using machine learning approaches. Four algorithms are implemented to the model, which are e Decision Tree, Logistic Regression, Naïve Bayes, and Random Forest algorithms [6]. The algorithms are compared based on the percentage of accuracy of test dataset. R and Python have been used for executing algorithms. Final algorithm is chosen with respect to the percentage accuracy and the false discovery rate which is obtained from each algorithm. As the conclusion, Logistic Regression is chosen as the he best algorithm for the model. The research determined the features that were the most significant for the prediction.

C. Diabetes Disease Prediction using Machine Learning on Big Data of Healthcare

This tells about model is suggested to predict diabetes disease using machine learning on big data of healthcare. This model uses WEKA tool to predict diabetes disease utilizing the Nave Bayes, Support Vector Machine, Random Forest, and Simple CART algorithms. The algorithm is selected based on various classifier measures and accuracy. As the results Support Vector Machine is chosen as the best fit algorithm to the model. The effectiveness of the proposed model is clearly depicted throughout the experimental results mentioned [7].

D. A Machine Learning Approach for Student Assessment in E-Learning using Quinlan's C4.5, Naïve Bayes, and Random Forest Algorithm.

This discuesess the model proposed to predict fair/transparent student evaluation using machine learning algorithms, which use different machine learning techniques to study students' performance. Classifiers like DecisionTrees-J48, Naïve Bayes and Random Forest were used to forecast the final results of the students based on the proposed model. The algorithms are evaluated based on the performance parameters that were selected and the prediction is made with each algorithm [8]. As the result proposed

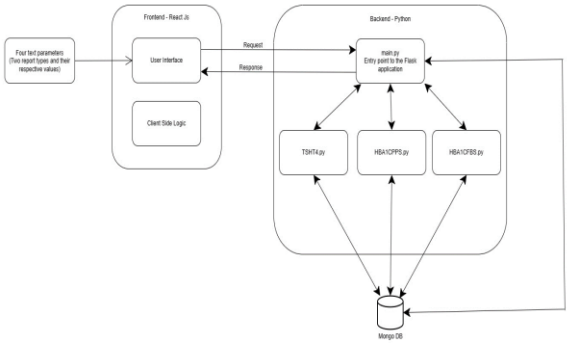


Fig. 1. Proposed System in a flowchart

model is very helpful in predicting students' results on behalf of their performance in prior tests and J48 algorithm shows the highest correctly classified instance percentage.

III. RESEARCH APPROACH

As the solution for to the problem, which is lack of health literacy in patients with related to medical test report results, a web based application is proposed to predict the medical test report results using machine learning algorithm and provide with information based on the prediction made. Four algorithms were used to find the best fit algorithm to the prediction model. The algorithm chosen are Naïve Bayes, Random Forest, Decision Tree, and Support Vector Machine algorithms. To develop the web application, React JS and Python are used as the frontend and backend languages respectively.

A. Data Acquisition Stage

The proposed system was implemented on five types of blood medical test report results. The prediction is made with respect to two blood medical test report results. The five blood medical test report results used are TSH, T4, HBA1C, FBS, and PPBS. The datasets are obtained from existing physical blood medical test reports and from Kaggle website. There are

three sets of two blood medical test report results dataset, which TSH and T4, HBA1C and FBS and HBA1C and PPBS. The attributes of the datasets are common to all, which are values obtained for each report and the expected result. The datasets are divided into training and testing data to the ration of 80%:20%

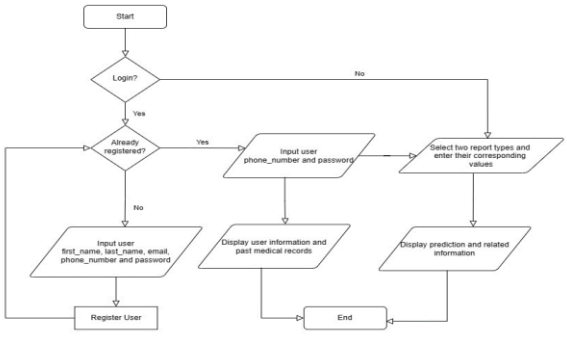


Fig. 2. System workflow of the developed web application

based on an evaluation conducted which showed the highest accuracy.

B. *Proposed Model and workflow of the proposed system.*

In the proposed system, it is designed to predict blood reports, and the system accepts two parameters, so that the prediction made is based on two medical test report results. The system is currently designed to five blood medical test reports, which are TSH, T4, HBA1C, FBC and PPBS, and there are three combinations to get the prediction, which are TSH and T4, HBA1C and FBC and HBA1C and PPBS. The system accepts four parameters to provide the prediction of medical test report results, which are report type 01, report type 01 value, report type 02, report type 02 value. Then based on the provided types and respective values the prediction is made and displayed to the user along with the detailed information about the prediction which includes what is meant by the prediction, risks, symptoms, precautions and two URL links which direct to a webpage that gives more information about the reports that the user has mentioned. The detailed information of the prediction was clarified and discussed with the assistance of a doctor. The system is also designed to allow users to register to the system and then login to the system later on. When the user logs in to the system, user can view user's information and the past medical test report results that includes the two report types the user entered and their

respective values, the prediction given and date and time that the prediction was made. If the user adds new medical test report result prediction, it will be also saved in the system and could be seen immediately in the past medical records. Above Figure 1 shows the proposed system in a flow chart diagram.

C. *Implementation*

The implementation of the proposed system is a web application. Web applications is chosen for the proposed system since it is easy to make updates in the system since the proposed system can be scaled to predict different report types and to give prediction based on more than two report types in the future. The web application is developed in React JS and Material UI framework for frontend developments and Python and Flask Framework for backend development.

There are three models developed in order to get prediction for TSH and T4 reports, HBA1C and FBC reports and HBA1C and PPBS reports. The models are developed in Python programming language. The algorithm used for prediction is Random Forest algorithm. Random Forest algorithm is chosen based on accuracy, which was compared with three other algorithms, which are Naïve Bayes, Decision Tree, and Support Vector Machine algorithms. The database used for storing and retrieving data is MongoDB. MongoDB is used for the system because it is flexible and can be easily expanded and also it can store images easily if need in the future.

The frontend of the system is developed using JavaScript library name Reat JS along with Material UI framework. Frontend of the system is used to get parameters from the user and display the required information to the user. The frontend and the backend of the system are communicated using APIs.

IV. Evaluation and Results

Algorithm for the models were chosen by evaluating four machine learning algorithm based on the accuracy of the training and testing data. The algorithms that were used are Naïve Bayes algorithm, Random Forest algorithm, Decision Tree algorithm and Support Vector Machine algorithm. The algorithms that were chosen to evaluate were taken based on the algorithm which was suggested and reviewed in the literature review chapter. The evaluation was conducted for each algorithm and the training data and testing data were varied. The percentage instances that the training data and testing data were taken are 20%: 80%, 25%: 75% and 30%: 70% respectively.

A. *Training data : Testing data – 80% : 20%*

TABLE I. Evaluation of algorithms based on the ratio training data: testing data - 80%:20%

	Gaussian Naïve Bayes	Random Forest	Decision Tree	Support Vector Machine
Training Data	86.50%	100.00%	100.00%	89.25%
Testing Data	89.50%	97.50%	96.50%	90.00%

B. *Training data: Testing data – 75%: 25%*

TABLE II. Evaluation of algorithms based on the ratio training data: testing data - 75%: 25%

	Gaussian Naïve Bayes	Random Forest	Decision Tree	Support Vector Machine
Training Data	87.20%	100.00%	100.00%	89.73%
Testing Data	88.40%	97.20%	96.80%	89.20%

C. *Training data: Testing data – 70%: 30%*

TABLE III. Evaluating of algorithms based on the ratio training data: testing data - 70%: 30%

	Gaussian Naïve Bayes	Random Forest	Decision Tree	Support Vector Machine
Training Data	89.29%	100.00%	100.00%	89.86%
Testing Data	88.00%	97.00%	96.00%	88.33%

After analyzing the results that were taken, which is mentioned in Table 1, Table 2 and Table 3, with highest accuracy in testing data and training data in all the three percentage ratios considered Random Forest Algorithm was chosen. Also, among the three percentage ratios that divide testing and training data, the ratio percentage of training and testing data that gives highest percentage of accuracy in Random Forest algorithm was chosen, which is 80%:20% respectively.

The final trained models were able to achieve an accuracy of 100% for training dataset and 97.50% for testing data with the ratio of training data and testing data as 80%:20%. The final system, which is a web application with machine learning algorithm, was as expected, and the requirements of the system were achieved.

V. Discussion and Conclusion

This research paper discusses a web application developed to address the research problem identified as the lack of health literacy among patients regarding medical test report results. The study began with a literature review, which identified similar systems and informed the workflow design of the web application, as well as the selection of the most suitable machine learning algorithm for the model of the system. The research approach proceeded in several stages, starting with data collection. This was followed by the proposal of a solution, including the identification of the best-fit algorithm for the system and the workflow of the system. In the final stage of the research approach, the proposed web application was implemented, focusing primarily on the development process. Lastly, the system was tested and evaluated to ensure that it met the expected results and requirements.

In conclusion, this research and the system development address a critical challenge faced by the patients when reading and understanding medical test report results without medical professional assistance. Use of machine learning algorithm, system offers the functionality of prediction and based on the prediction the relevant information a given to the user. The developed web application meets the requirements, and the results expected. This system caters to both technical and non-technical users and both medical and non-medical professionals emphasizing user-friendly interfaces. The system also offers to see the past medical results which are previously entered by the user, which helps the user to get an idea about the history of their results. This study is a significant step toward lack of medical literacy in patients on medical test report results

VI. Future Recommendations

Future Recommendations includes expanding the current development to provide prediction and summarized information for more blood reports and for various report types such as urine, saliva, and solid waste. The system can be improved to upload images of the reports as an input method and to get more than two medical test report results for prediction from the user. In the future development, the system can be developed to see the past medical report results analyzed and displayed using a data representation method like bar chat, pie chart. When expanding the system, the system can be developed in micro-frontend architecture which frontend can be divided as blood, urine, solid waste, so that the developers can expand whenever there's a requirement.

Mentioned future recommendations aim to improve the system's analysis and prediction skills, expand the system to predict and gain knowledge about more medical reports and incorporate user-centric elements for a more holistic and user-friendly experience.

Acknowledgment

I'd like to take this opportunity to thank everyone who has helped me along the way with my research. Initially, I would like to pay my gratitude to Mr.Pramudya Thilakarathne, for providing me with unwavering support, guidance and expertise thought out this research project. All his insightful advice and encouragement have been invaluable. I would like to express my deep appreciation to Dr. A.C.V. Silva for her assistance in educating me on health matters, particularly in understanding test reports and their implications. Next, I would like to extend my heartfelt thanks to my parents and family for their unconditional and constant support and understanding throughout this journey. I would like to express my heartful thanks to the medical professionals and other individuals who helped me to gather information for this research. Their real-world insights into the Healthcare Industry regarding medical test report results proved to be extremely helpful in shaping the direction of this study.

References

[1] N. D. Berkman, T. C. Davis, and L. McCormack, "Health literacy: What is it?," 2010. doi: 10.1080/10810730.2010.499985.

[2] P. N. Peterson et al., "Health Literacy and Outcomes Among Patients With Heart Failure." [Online]. Available: https://jamanetwork.com/

[3] Lynn. Nielsen-Bohlman, A. M. Panzer, D. A. Kindig, and Institute of Medicine (U.S.). Committee on Health Literacy., Health literacy : a prescription to end confusion. National Academies Press, 2004.

[4] S. S. Coughlin, M. Vernon, C. Hatzigeorgiou, and V. George, "Health Literacy, Social Determinants of Health, and Disease Prevention and Control."

[5] Sri Shakthi Institute of Engineering and Technology, Institute of Electrical and Electronics Engineers. Madras Section, All-India Council for Technical Education, and Institute of Electrical and Electronics Engineers, 2020 International Conference on Computer Communication and Informatics : January 22-24, 2020, Coimbatore, India.

[6] P. N. Astya, Galgotias University. School of Computing Science and Engineering, Institute of Electrical and Electronics Engineers. Uttar Pradesh Section, and Institute of Electrical and Electronics Engineers, IEEE International Conference on Computing, Communication and Automation (ICCCA 2017) : proceeding : on 5th-6th May, 2017.

[7] 2018 Fourth International Conference on Computing Communication Control and Automation (ICCUBEA).

[8] P. Air University (Islamabad and Institute of Electrical and Electronics Engineers, Proceedings of the 2016 19th

International Multi-Topic Conference (INMIC) : 05-06 December 2016, Faculty of Electrical Engineering, Air University, Islamabad, Pakistan.

Cricketscorenet: Machine Learning-based Cricket Score Prediction App for Sri Lanka T20 with Advanced Features

KAHK Karandana
Department of Computer Science
General Sir John Kotelawala Defence University
Ratmalana, Sri Lanka
38-bcs-0014@kdu.ac.lk

Mrs. Mihiri Sirisuriya
Department of Computer Science
General Sir John Kotelawala Defence University
Ratmalana, Sri Lanka
mihiri@kdu.ac.lk

***Abstract*— Cricket is a sport of skill, strategy, and adaptability. It is deeply influenced by the dynamic interplay of numerous factors. Making the prediction of match scores is a complex challenge.Over the past few years, machine learning has made this better use of predictive outcomes for sports more insightful. Specially, Cricket uses machine learning. In this paper, we present an end-to-end T20 cricket match result prediction solution. This study presents a machine learning approach for predicting outcomes of T20 cricket matches in Sri Lanka. We developed models to forecast the winning team pre-match, batting team scores during innings, and second innings winners. It is interrogated with conventional statistical methods along with contemporary machine learning algorithms. [1] Our approach not only improves current prediction models by adding other pertinent disregarded factors like pitch variations, player injuries or weather but also analyses the possible influence that some of these variables can have in other ones. We use extensive image analysis of pitch images with a Convolutional Neural Network. It makes our predictions far more precise and detailed. The study is dedicated to predicting the winning team in a T20 cricket match played at Sri Lanka. [2]Therefore, it produces a fast prediction tool for efficient results and supports decision-making for teams and analysts. Our approach represents a significant step forward in cricket analytics by offering a powerful resource for understanding and predicting the complex dynamics of T20 cricket matches.**

Keywords*— *Match Score Prediction, Machine Learning, Ball Trajectory Prediction, Sentiment Analysis, Sports Analytics

I. INTRODUCTION

Easily one of the most established global games, Cricket has an unpredictable style which is relied upon to deliver a great deal of factors .It normally applies when it arrives into foreseeing outcomes. The fast and unpredictable pace of the T20 format does not make this any easier. The only accurate way to predict the outcome of cricket matches is with historical data analysis which has always been used as a traditional method. [3] These methodologies are often fairly limited in accuracy due to the absence of dynamic and contextual factors such as pitch conditions, player injuries or weather, Although Those are extremely insightful.

The recent trends in machine learning have enabled a completely new paradigm for sports analytics . It offers much more sophisticated and accurate prediction models. In this paper, we delve into a set of machine learning algorithms for enhancing the predictive capability on cricket outcomes. They leverage smaller patterns in data and use it. And we make predictions at a minimal scale. Existing researches have mostly focused on One Day Internationals (ODIs) or Test matches but hardly any dedicated to T20s. In addition, existing models usually do not feature real-time data inputs . So, this is necessary for a model that can produce dynamic and accurate predictions.

Our research attempts to close these gaps by creating an extensive prediction system especially suited for T20 cricket matches in Sri Lanka. Our method combines cutting-edge machine learning algorithms with conventional statistical techniques to produce a reliable prediction tool. We improve our models by adding variables like weather, player injuries, and pitch conditions those are causing a big influence on match results. But Those are frequently left out of previous models.

The use of Convolutional Neural Networks for pitch picture analysis is a significant novelty in our methodology. The state of the pitch has a significant impact on player performance, scoring rates, ball behavior, and other match-related variables. A more sophisticated knowledge of pitch conditions and their effects on the game is possible by using CNNs to extract specific elements from pitch photos.

Here is the breakdown of our study. We start with data gathering and preprocessing to ensure the comprehensiveness and balance of our dataset. Subsequently, we create three discrete models to forecast several parts of the cricket match .The team that wins before the game begins, the batting team's score throughout an innings, and the team that wins in the second innings.

Our prediction tool can be accessed in real-time by users through an application that is designed and implemented on the web. With the use of precise and current forecasts, this program is meant to assist analysts and teams in making well-informed choices.

This research intends to construct a dynamic and highly accurate prediction tool for T20 cricket matches by integrating several predictive criteria and utilizing modern machine learning algorithms. Our objective is to give teams and analysts a stronger tool for comprehending and predicting match dynamics, so improving their strategic intelligence. Offering fresh perspectives and resources to enhance performance and decision-making in Twenty20 cricket, this study is a noteworthy development in the field of cricket analytics.

II. Related Works

The application of machine learning to sports analytics is growing in popularity, providing new ways to comprehend and more accurately forecast sports results. Predictive modeling is both challenged and benefited by the intricacy of the cricket game, which is characterized by a multitude of influencing elements such player performance, match conditions, and strategies. With an emphasis on data mining, statistical analysis, and the use of artificial intelligence, the literature on cricket match prediction has investigated a variety of machine learning methods and methodologies.

In order to forecast the results of cricket matches, Naik et al. investigated the applications of logistic regression, neural networks, and K-means clustering. The study emphasized the difficulties in gathering and interpreting data and the need of choosing suitable features and algorithms to enable precise prediction-making. The quality and completeness of the dataset had a significant impact on the neural networks' ability to detect intricate patterns in the data, the researchers discovered.

For the purpose of conducting experimental verifications and making well-informed business decisions, Jyothsna and Srikanth concentrated on the application of data analytics in sports, specifically cricket. The importance of applying sophisticated analytics to understand player performance, team tactics, and match results was highlighted by their study. They underlined that in order to increase forecast accuracy, thorough data gathering and the use of advanced analytical tools are essential.

Artificial intelligence algorithms for forecasting the results of Test and ODI matches were the subject of a 2018 SADP study. The study brought attention to the shortcomings of the previous models, which frequently failed to take into consideration how dynamic T20 matches are. With T20 cricket's particular challenges and quick tempo in mind, the writers demanded the creation of comparable instruments.

After reviewing the literature, the following important conclusions and gaps are found.

Prioritize ODI and Test Matches- The majority of research has focused on forecasting results for ODI and Test matches. The T20 format, which has unique properties and necessitates alternative modeling methodologies, has been the subject of relatively little research.

Differentiating between Static and Dynamic forecasts-Current models usually offer static forecasts that are derived from pre-match or sparse in-game data. Due to the fast-paced nature of Twenty20 cricket, there is a dearth of technologies that provide dynamic, real-time forecasts during the contest.

Restricted Feature Sets- A lot of models lack a full range of characteristics, which are known to have a big impact on match results. Examples of these features include weather, player injuries, and pitch conditions.

Algorithmic approaches- Although a number of algorithms,

such as decision trees, neural networks, and logistic regression, have been investigated, more complex models remain needed to manage the intricacy and unpredictability of cricket matches.

The body of study on cricket match prediction that has already been done offers insightful information and a starting point for future studies. To increase prediction models' application and accuracy, there are, nevertheless, several glaring gaps that must be filled, especially with regard to T20 cricket. By creating an all-encompassing prediction system that uses cutting-edge machine learning methods, such as Convolutional Neural Networks (CNNs) for pitch picture processing, and integrates extra predictive variables, this research seeks to close these gaps. Our goal in doing this is to produce a dynamic, real-time prediction tool that improves T20 cricket teams' and analysts' strategic capabilities

III. Methodology

The system that has been suggested for forecasting the results of Twenty20 cricket matches combines conventional statistical methods with advanced machine learning algorithms. It is further refined by adding variables like weather, player injuries, and pitch conditions. The overall goal of this thorough strategy is to greatly increase the prediction accuracy of match outcomes.

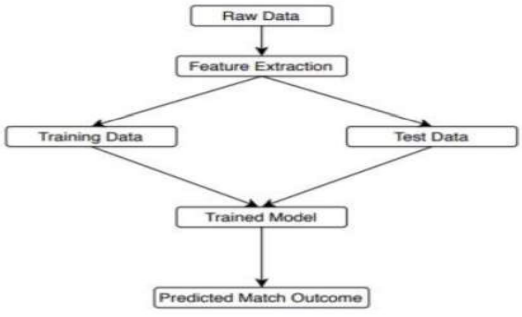


Fig. 1. *Overview of the Project*

A. Gathering and preprocessing data

In the first step, a lot of data is gathered from several Twenty20 cricket competitions, such as the Big Bash League, T20 International, IPL, and domestic T20 leagues in Sri Lanka. The collection includes ball-by-ball data, player statistics, match facts, [4]and historical weather information. Synthetic Minority Over-sampling Technique (SMOTE) is one of the oversampling strategies used to ensure an even distribution of classes in order to solve the imbalance in the dataset. There are also fresh data sources included:

Pitch Conditions:-Pictures of the pitches are taken prior to every game.

Player Injuries- Data regarding player injuries and how they affect performance is gathered.

Weather - Each match location's historical and current weather information is gathered.

Cleaning up the dataset through the removal of duplicates, correction of inconsistencies, and handling of missing values is known as data preparation. To generate new features from raw data, such as average player performance metrics, pitch conditions taken from photos, and weather condition indices, feature engineering is used. To boost variability and strengthen model robustness, pitch images go through preprocessing procedures such shrinking, standardizing pixel values, and data augmentation.

B. Model Construction

Three models are created to forecast various elements of the cricket match.

Forecasting the Winning Team Before the Match-Random Forest algorithm inputs include team names, venue, toss outcome, and decision.

Output - The likelihood of each squad winning

A score Forecast for the Squad Batting in an Inning -Ball-by-ball data, such as the batting team, bowling team, striker, non-striker, bowler, runs, additional runs, and wickets, are entered.

Modeling Hybrid Teams to Win in the Second Innings -Combination of Decision Tree and Random Forest Algorithm Inputs: Ball-by-ball information from the second inning as well as outputs from Models 1 and 2.

Output - The likelihood of each squad winning

In order to improve forecast accuracy, the models incorporate more variables including weather, player ailments, and pitch characteristics. Pitch picture analysis using Convolutional Neural Networks (CNNs) is a significant innovation. The state of the pitch has a big impact on player performance, scoring rates, ball behavior, and match dynamics. A more sophisticated understanding of pitch conditions and their effects on the game is made possible by the use of CNNs to extract detailed information from pitch photos.

With player injuries and weather included as extra elements in all models, a more thorough examination of the variables affecting match results is possible.

C. Design and Implementation of Systems

The real-time prediction, model training, and data preparation modules make up the system architecture. An intuitive online application is created [2]that allows users to enter match information and get real-time predictions. With the use of precise and current forecasts, teams and analysts can make well-informed judgments with the help of this program. The objective of incorporating many predictive criteria and sophisticated machine learning algorithms is to generate a resilient and adaptable prediction tool that will augment the strategic abilities of T20 cricket teams and analysts. Figure 01 represents the design overview of the system.

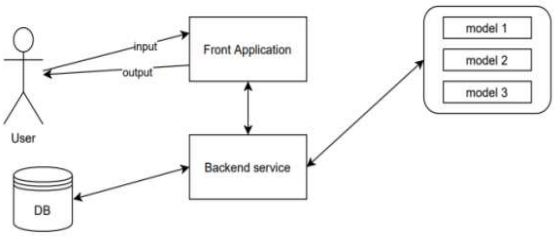


Fig. 2. Overview design of the system

IV. Results And Discussion

Data from a number of Twenty20 cricket competitions, including the Big Bash League, T20 International, IPL, and domestic T20 leagues in Sri Lanka, were used to test the suggested methodology. The prediction accuracy of the models was assessed while accounting for the incorporation of other data including weather, player injuries, and pitch conditions.

A. Predicting the Winning Team Before the Game Begins

84.51% accuracy on unbalanced data

76.61% accuracy on balanced data

Prior to the game, the Random Forest model demonstrated a high degree of accuracy in predicting which team would win. The model performed better when the data was uneven, suggesting that match [4] outcomes' inherent imbalance may include significant predictive information. But the balanced data technique ensures robustness across many datasets and offers a more generalized model.

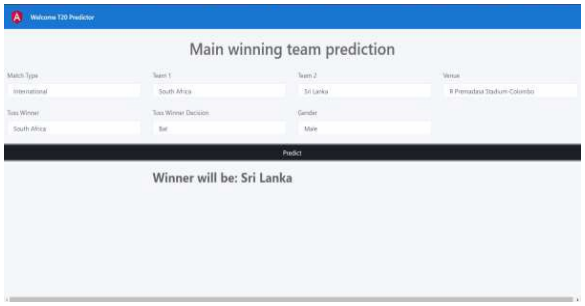


Fig. 3. Preview of Model 01

B. Predicting the Team's Score During an Inning

Decision Tree Algorithm Mean Absolute Error: The model's low mean absolute error suggested that the scores were accurately predicted. The efficiency of this model is critical to comprehending game dynamics and possible targets, particularly in the first innings.

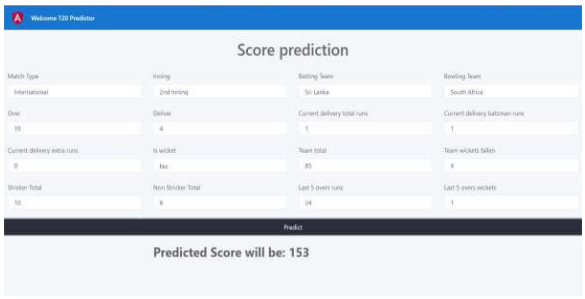


Fig. 4. Preview of Model 02

C. Hybrid Approach for Predicting the Winning Team in the Second Innings

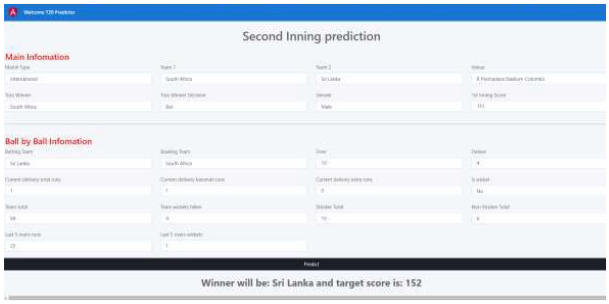


Fig. 5. Preview of Model 02

Algorithm - Random Forest + Decision Tree Combination;

97.23% Accuracy on Unbalanced Data

93.92% Accuracy on balanced data

Particularly when dealing with unbalanced data, the hybrid model which combines the advantages of the Random Forest and Decision Tree algorithms displayed remarkable accuracy. This impressive level of success highlights the model's flexibility in adjusting to the second innings' dynamic conditions by utilizing both previous models' forecasts and real-time ball-by-ball data.

The models' predicted accuracy was greatly increased by the addition of new factors. Pitch conditions were included, and Convolutional Neural Networks (CNNs) analysis of the data produced important insights that traditional data could have missed. Since ball behavior and player performance are known to be influenced by pitch conditions, the ability to measure both via image analysis is a major improvement.

Effect of Pitch Conditions- Pitch images were successfully used by the CNN model to extract features, which improved prediction accuracy. With this method, the system could take into account the subtle impacts of pitch wear and features on the results of matches. This creative implementation of image analysis in sports analytics shows how applicable it could be in other fields.

Impact of Player Injuries and Weather: Including information on player injuries and weather gave the prediction models an additional level of complexity. A team's ability to compete can be greatly impacted by injuries, and the context of current weather data can alter the dynamics of a game. The inclusion of these extra characteristics improved the models' comprehensiveness and increased their realism.

Performance of the Models: Combining several algorithms and data sources is important, as evidenced by the hybrid model's high accuracy in forecasting the winning team in the second innings. Through the combined use of Random Forest and Decision Tree algorithms, the hybrid model was able to efficiently manage the complexity of real-time predictions.

Real-Time Prediction Capabilities: This methodology's web-based application turned out to be effective and user-friendly in generating real-time forecasts. This useful tool demonstrates how machine learning may improve sports analytics by helping teams and analysts make critical decisions during a game.

V. Future Works

Even with the achievement, a number of obstacles still exist. Model performance may be impacted by the unpredictability of data quality, especially for recent and poorly reported tournaments. In the future, research will concentrate on enhancing data gathering techniques and investigating new elements like player shape, psychological aspects, and sophisticated weather predicting models. More complex neural network topologies and more CNN model refinement to handle a wider range of pitch situations may also result in predictions that are even more accurate.

VI. Conclusion

In order to fill important gaps in the current models, this study offers a thorough methodology for forecasting the results of Twenty20 cricket matches by combining conventional statistical methods with modern machine learning algorithms. Our method offers a more detailed view of match dynamics by integrating dynamic aspects like pitch conditions, player conditions, and weather, and by using neural network models (CNNs) for pitch picture interpretation. The hybrid model, which included Decision Trees and Random Forest, showed remarkable accuracy, especially when it came to second innings forecasts. The web-based technology designed for in-the-moment forecasts has shown to be successful, providing teams and analysts with an invaluable resource to help them make judgments during games. In order to make even more accurate predictions, future research will concentrate on enhancing data quality, investigating new variables like player form and psychological aspects, and optimizing neural network topologies. This study represents a substantial breakthrough in cricket analytics, strengthening T20 teams' strategic capacities and

promoting improved decision-making and performance gains in the game as a whole.

ACKNOWLEDGMENT

I would like to express my sincere gratitude to my supervisor Mrs. Mihiri Sirisuriya, who supported me throughout this research project with utmost dedication and motivating me. I would like to express my gratitude to all the lecturer staff in Department of Computer Science, General Sir John Kotelawala Defence University, parents and family by encouraging and supporting me throughout this research. As this is a collabarative effort, I'm thankful for all the contributions given for the success of this research paper.

REFERENCES

[1] S. Y. A. G. M. I. H. a. A. C. A. I. Anik, "Player's Performance Prediction in ODI Cricket Using Machine Learning Algorithms," 2018.

[2] Y. &. R. T. Wang, "Predictive Modeling of Cricket Match Outcomes Using Machine Learning Algorithms.," 2019.

[3] J. &. J. L. Smith, "The Impact of Weather on Cricket Match Outcomes," 2018.

[4] A. &. G. T. Foster, "Advanced Analytics for Cricket Performance," 2017.

[5] N. K. K. A. A. D. R. P. Prof. R. R. Kamble1, Cricket Score Prediction Using Machine Learning, 2021.

[6] A.A.Kaluarachchi, ""CRICAI a classification based tool for odi match outcome prediction"," 2010.

[7] P. a. Hatharasinghe, "data mining and machine learning in cricket match outcome prediction," 2019.

[8] M.Kowsher, "Detecting Third Umpire Decisions Automated Scoring System of Cricket," 2019.

[9] N. V. K. M. V. N. N. S. T. Akhil Nimmagadda1, "Cricket score and winning prediction using data".

[10] P. R. P. K. ,. N. K. Preetham HK, "Cricket Score Prediction Using Machine Learning," 2023.

[11] R. D. A.N.Wickramasinghe, "Cricket Match Outcome Prediction Using Tweets and Prediction of the Man of the Match using Social Network," 2018.

[12] N. S. S. R. V. S. Nigel Rodrigues, "Cricket Squad Analysis Using Multiple Random Forest Regression," 2019.

[13] S. K. a. S. Roy, "Score Prediction and Player Classification Model in the Game of Cricket Using Machine Learning," 2018.

[14] A. &. S. S. Kumar, "Advancements in Sports Analytics," 2020.

[15] K. R.watson, " Machine Learning Techniques for Predicting Cricket Match Scores," 2019.

[16] P. &. S. M. Jones, "Ball Tracking Technology and Its Impact on Cricket Analysis," 2017.

[17] M. B. a. S. R. Clarke, PREDICTING THE MATCH OUTCOME IN ONE DAY INTERNATIONAL CRICKET MATCHES WILL THE GAME IS IN PROGRESS, 2006.

[18] M. ,. A. A. M. J. U. Kowsher, "Detecting Third Umpire Decisions & automated scoring system of cricket.," 2019.

[19] A. R. P. o. C. P. U. V. M. L. A. a. c. a. them, "A Review Paper on Cricket Predictions Using machine learning".

Impact of Generative AI on the Use and Reliability of Video Evidence

Diluka Wijesinghe
Dept. of Software Engineering & Information Systems
NSBM Green University
Mahenwaththa, Pitipana, Homagama, Sri Lanka
diluka.w@nsbm.ac.lk

***Abstract*— The emergence of generative artificial intelligence (AI), particularly in the form of deepfakes, presents profound challenges to the reliability and admissibility of video evidence within legal and investigative frameworks. This paper critically examines the implications of AI-generated video manipulation on judicial processes, focusing on the erosion of evidentiary trust and the inadequacies of existing detection methodologies. Through a rigorous analysis of state-of-the-art detection techniques and the evolving legal landscape, this study underscores the urgent need for technological, legal, and ethical advancements to counter the risks posed by AI-manipulated media. Key findings indicate that current forensic tools are insufficient in identifying sophisticated deepfake content, necessitating the development of more robust verification mechanisms and a reevaluation of evidentiary standards. The paper concludes by proposing a multidisciplinary approach to fortify the integrity of video evidence in the AI era, combining technological innovation, legal reform, and international regulatory collaboration.**

Keywords*— *Generative AI, Deepfakes, Video Forensics, Legal Standards, AI Ethics, Digital Manipulation

I. INTRODUCTION

The advent of generative artificial intelligence (AI) and its application in creating deepfakes has introduced unprecedented challenges to the reliability and authenticity of video evidence. Deepfakes, which are hyper-realistic synthetic media generated using techniques such as generative adversarial networks (GANs), have the potential to convincingly alter video and audio content, making it difficult to distinguish between real and fabricated footage. This technology has sparked significant concerns in legal, ethical, and societal domains due to its ability to manipulate reality with high precision.

Video evidence has traditionally been one of the most trusted forms of evidence in legal proceedings due to its visual and auditory components, which offer seemingly irrefutable proof of events. However, the introduction of deepfakes threatens this trust, as these AI-generated videos can be used to create deceptive content that appears authentic. This poses a severe threat to the integrity of legal processes, where the authenticity of evidence is paramount [1].

The implications of deepfakes extend beyond the courtroom. They have been employed in various malicious activities, such as disinformation campaigns, political manipulation, and personal vendettas, thereby eroding public trust in digital media. As deepfake technology becomes more accessible and sophisticated, the potential for misuse increases, raising critical ethical concerns and necessitating new legal frameworks to address these challenges [2][3].

The legal system faces a significant challenge in adapting to this new reality. Current standards for the admissibility and authentication of video evidence are not fully equipped to handle the complexities introduced by AI-generated content. The burden of proof, traditionally placed on the party presenting the evidence, becomes complicated when the authenticity of video content can be easily questioned due to the potential for deepfakes [2].

II. LITERATURE REVIEW

A. Introduction to Generative AI

Generative AI encompasses advanced machine learning models designed to create new data instances that mimic a given dataset. Among the primary frameworks are Generative Adversarial Networks (GANs) and Variational Autoencoders (VAEs). GANs, introduced by Goodfellow et al. [4], involve a dual-network system: a generator and a discriminator that engage in adversarial training to enhance the realism of generated data. VAEs, proposed by Kingma and Welling [4], utilize probabilistic models to learn latent variables and generate new samples from learned distributions. These technologies have transformative applications, including image synthesis, creative arts, and video production, raising significant implications for the integrity and trustworthiness of video evidence.

B. Generative AI and Video Evidence

Generative AI technologies, particularly deepfakes, have significantly altered the landscape of video evidence. Deepfakes utilize GANs to create highly realistic video content by manipulating facial features, voices, and expressions [5]. This capability has profound implications for video evidence, enabling the creation of convincingly fabricated content that challenges traditional verification processes [6].

Techniques such as face-swapping and audio synchronization have demonstrated the potential to produce deceptively realistic videos, complicating the verification process for legal and investigative purposes.

C. Impact on the Use of Video Evidence

Detection techniques for manipulated video content include forensic algorithms and deep learning models designed to identify inconsistencies and anomalies in video data [7]. Advances in AI forensics focus on improving the accuracy and efficiency of these detection methods [8]. Traditional verification methods are increasingly inadequate as generative technologies evolve, highlighting the need for new forensic tools and methodologies [9].

D. Reliability Concerns

The rise of generative AI technologies raises significant concerns about the reliability of video evidence. Traditional verification methods struggle to keep pace with the realism of generated content, necessitating the development of new forensic approaches [9]. Additionally, the potential for misinformation is exacerbated by the ability of generative AI to produce realistic yet false videos, which can distort public opinion and contribute to societal harm [10]. Ethical considerations, including privacy, consent, and the potential for misuse, are crucial in addressing the responsible use of generative AI for video manipulation [11].

E. Countermeasures and Detection

To combat the challenges posed by generative AI, various detection methods are being refined, including forensic algorithms and deep learning models [12]. Advances in AI forensics aim to enhance the effectiveness of these techniques, addressing the growing sophistication of deepfake technologies [13]. Effective policy and regulation are essential to manage the risks associated with generative AI, including developing guidelines for responsible use and controlling the dissemination of manipulated content [4].

F. Case Studies and Real-World Examples

High-profile cases involving deepfakes, such as political misinformation campaigns and celebrity defamation, illustrate the practical implications of manipulated video evidence [14]. Analyzing these cases reveals broader implications for legal proceedings, public trust, and law enforcement's ability to manage synthetic media [9].

G. Future Directions

As generative AI technologies continue to advance, they will present new challenges for video evidence. Researchers should anticipate these developments and prepare adaptive strategies for detection and verification [8]. Future research should focus on developing more robust detection methods, enhancing forensic techniques, and implementing comprehensive policies to mitigate the risks associated with generative AI [11].

Although numerous studies have explored the detection of deepfakes, there are significant gaps in the literature. Most notably, existing research lacks a comprehensive evaluation of the scalability of detection tools as deepfake technology becomes more accessible. Additionally, few studies examine the legal implications of AI-generated video evidence in courtrooms, where current standards are insufficient to handle the complexity of AI-manipulated media.

III. METHODOLOGY

This study utilizes a comprehensive methodological approach to examine the impact of generative AI on video evidence. The approach integrates theoretical analysis, empirical case studies, and technical evaluation to provide a detailed understanding of how generative AI technologies influence the credibility and use of video evidence.

A. Theoretical Analysis

The study begins with a theoretical analysis of generative AI technologies, specifically Generative Adversarial Networks (GANs) and Variational Autoencoders (VAEs). This involves a thorough review of existing literature to comprehend the fundamental principles, architectures, and training processes of GANs and VAEs. The analysis covers seminal research papers, technical reports, and recent advancements to grasp the capabilities of these models in generating and manipulating video content.

B. Empirical Case Studies

To complement the theoretical analysis, the study conducts empirical case studies of real-world instances where generative AI technologies have been applied to video manipulation. Case studies are selected based on their relevance to the use of generative AI in video creation and their impact on social, legal, or ethical dimensions. The analysis involves reviewing news reports, legal documents, and academic articles to assess how generative AI was employed, the techniques used, and the resulting implications for video evidence integrity.

C. Technical Evaluation

The technical evaluation focuses on assessing the effectiveness of various detection methods for manipulated video content. This includes a review of existing detection techniques, such as forensic algorithms and machine learning models, as well as emerging technologies like blockchain for video verification. The evaluation involves comparing these methods based on performance metrics such as accuracy, false positives, and false negatives, and identifying strengths and weaknesses in current detection technologies.

D. Data Sources

Data for the study is collected from multiple sources to support the theoretical analysis, case studies, and technical evaluation. Key sources include:

- Literature: Peer-reviewed academic journals, technical reports from research institutions, and conference papers discussing

advancements in generative AI and video manipulation.

- Case Studies: Media coverage, legal documents, and research papers detailing incidents of video manipulation and its impact.
- Detection Techniques: Research on forensic analysis and machine learning-based detection methods, industry reports on advancements in video detection technologies, and documentation for relevant detection tools.

E. Analytical Framework

The analysis employs a robust framework combining theoretical, forensic, and legal perspectives. This framework includes:

- Generative AI Theory: Examining the mechanisms and capabilities of GANs and VAEs in generating and altering video content.
- Forensic Science Principles: Applying forensic principles to evaluate the authenticity and integrity of video evidence.
- Legal and Ethical Considerations: Incorporating legal and ethical perspectives to assess the implications of video manipulation and the effectiveness of current detection methods.

Additionally, the study reviews various generative AI models and detection methods. It includes an analysis of GAN architecture and training processes, VAEs for video synthesis, machine learning models for detecting anomalies, and forensic techniques for identifying manipulation traces. A comparative analysis benchmarks detection methods and evaluates their performance in real-world scenarios.

IV. Generative AI Technologies

Generative AI has significantly advanced the field of artificial intelligence by enabling the creation of highly realistic synthetic data. This section examines the foundational mechanisms of generative AI, focusing on Generative Adversarial Networks (GANs) and Variational Autoencoders (VAEs), and explores their applications in video generation and manipulation.

A. Generative Adversarial Networks (GANs)

Generative Adversarial Networks (GANs), introduced by Goodfellow et al. [4], consist of two neural networks—the generator and the discriminator—that engage in a competitive process. The generator aims to create synthetic data from random noise, while the discriminator evaluates the authenticity of this data compared to real data. This iterative process continues until the generator produces data that the discriminator cannot distinguish from genuine data, resulting in highly realistic synthetic outputs.

B. Variational Autoencoders (VAEs)

Variational Autoencoders (VAEs) apply a probabilistic approach to data generation. They consist of an encoder and a decoder. The encoder maps input data to a latent space, representing it as a distribution rather than a fixed point, capturing the underlying variations. The decoder reconstructs data from this latent space, generating new instances by sampling from the learned distribution. VAEs balance reconstruction accuracy with latent space compactness, facilitating the generation of diverse and coherent data.

C. Techniques for Video Generation and Manipulation

Generative AI techniques, particularly deepfakes, have transformed video content creation and manipulation. Deepfakes leverage GANs to produce hyper-realistic videos by altering or fabricating the appearance and actions of individuals. Key techniques include:

- **Face Swapping**: Replacing an individual's face in a video with another's, achieved through GANs trained on extensive facial datasets.
- **Lip Syncing**: Synchronizing lip movements with an audio track to create convincing fake speech videos.
- **Motion Transfer**: Transferring the movements of one person to another, enabling the creation of videos depicting actions not originally performed by the target individual.

These techniques exploit GANs' capability to generate realistic facial expressions and movements, posing significant challenges in distinguishing real from manipulated video content.

V. Manipulation of Video Evidence

The advancement of generative AI has introduced sophisticated methods for creating and manipulating video content, leading to significant concerns about the authenticity and reliability of video evidence. This section explores the techniques used for generating deepfakes, methods for detecting them, and provides real-world examples to illustrate these manipulations.

A. Creating Deepfakes

Face Swapping Face swapping involves using Generative Adversarial Networks (GANs) to replace the face of one individual with another in a video. This technique is facilitated by training models on images of both the target and source faces, enabling accurate mapping of facial expressions and movements. While face swapping has applications in entertainment and media, it also raises risks related to misinformation and fraud.

Lip Syncing Lip syncing generates videos where the lip movements of a person align with an audio track. This is achieved through audio-driven GANs or Recurrent Neural Networks (RNNs), which synchronize lip movements with spoken words. This

technique is used to create realistic speech videos for dubbed content and virtual avatars, enhancing the realism of digital media.

Motion Transfer Motion transfer involves animating one individual's movements onto another. Techniques such as the First Order Motion Model utilize keypoint detection and motion fields to animate the target subject. This method is employed to create animated versions of still images or historical figures and for virtual puppetry.

B. Detecting Deepfakes

Artifact Detection Artifact detection focuses on identifying inconsistencies and anomalies in videos, such as unnatural eye movements, inconsistent lighting, or irregular facial features. Tools like DeepFake-o-meter and FaceForensics++ leverage machine learning to detect these subtle artifacts.

Deep Learning-based Detection Deep learning-based detection employs Convolutional Neural Networks (CNNs) and other advanced architectures to identify deepfakes. These methods involve training models on extensive datasets of real and fake videos to recognize patterns indicative of manipulation.

Temporal Analysis Temporal analysis examines the coherence of video frames over time to identify irregularities. This method studies the consistency of facial expressions and movements frame-by-frame to detect anomalies that suggest manipulation.

Watermarking and Digital Signatures Watermarking and digital signatures embed invisible markers into video content for authentication. Cryptographic techniques are used to ensure that any alteration can be detected by verifying the integrity of the embedded watermark or digital signature.

VI. Impact on Legal and Judicial Systems

Generative AI technologies have profoundly impacted the legal landscape, introducing both opportunities and significant challenges regarding video evidence. This section explores the legal ramifications, the difficulties in verifying video authenticity, and the implications for law enforcement and investigative processes.

A. Legal Implications of AI-Generated Video Evidence

The rise of AI-generated video content necessitates the establishment of new evidentiary standards. Courts are required to develop robust criteria for assessing the authenticity and admissibility of digital evidence. This involves ensuring a secure chain of custody to prevent tampering and guarantee the integrity of the evidence.

AI-generated videos pose risks of fabrication, which can lead to wrongful accusations or defenses. Legal frameworks must address the complexities of liability associated with the creation and dissemination of deepfakes, encompassing defamation and privacy violations. As a response, jurisdictions are beginning to implement targeted legislation that criminalizes harmful uses of AI-generated content, alongside developing comprehensive regulatory frameworks to govern the ethical use of these technologies.

B. Challenges Faced by the Judicial System in Verifying Video Authenticity

The judicial system faces significant challenges in verifying the authenticity of video evidence due to the rapid evolution of AI technologies. Expert testimony has become increasingly crucial, requiring courts to rely on forensic analysts with specialized skills in detecting AI-manipulated content. Additionally, the judicial system must continuously update its knowledge and capabilities to keep pace with technological advancements.

Resource constraints further complicate the verification process. Analyzing and validating digital evidence can be both time-consuming and costly, straining judicial resources. Smaller jurisdictions, in particular, may struggle with limited access to advanced forensic tools necessary for effective deepfake detection. The traditional presumptions of video authenticity are also challenged by sophisticated manipulation techniques, requiring new methods for establishing and proving the legitimacy of video content.

C. Impact on Law Enforcement and Investigations

Generative AI technologies have significant implications for law enforcement and investigative processes. The potential for evidence tampering through deepfake technology complicates investigations and increases the risk of wrongful convictions or acquittals. Additionally, the spread of misinformation through fake videos can mislead investigations and divert resources from genuine leads.

To address these challenges, law enforcement agencies must invest in both training and advanced technology to stay abreast of developments in generative AI. Effective strategies require enhanced collaboration with technology companies, cybersecurity experts, and academic researchers. Public trust in the justice system depends on the credibility of video evidence, making it essential for law enforcement to engage in public awareness campaigns about the risks of deepfake technology and promote digital literacy.

VII. Case Studies

A. The "Nancy Pelosi Deepfake"

In 2019, a manipulated video of U.S. Speaker Nancy Pelosi was disseminated on social media, depicting her as inebriated. This deepfake was created using DeepFaceLab, which leverages Generative Adversarial Networks (GANs) to overlay altered facial expressions on original footage. The GAN was trained on a dataset of Pelosi's images to generate synthetic expressions and lip movements. Detection involved forensic analysis combined with machine learning techniques to identify discrepancies in audio-visual synchronization and facial motion artifacts.

B. The "BBC Deepfake Experiment"

In 2020, the BBC conducted an experiment to demonstrate the capabilities and potential misuse of deepfake technology by fabricating an interview with a prominent public figure. This deepfake was produced using advanced Variational Autoencoders (VAEs) and voice synthesis algorithms, which simulated realistic facial features and speech patterns. Detection methods included analyzing video metadata and employing AI-based forensic tools to identify inconsistencies in the audio-visual data.

C. "Deepfake Pornography Scandals"

High-profile incidents involving deepfake pornography have emerged, where celebrities' faces were superimposed onto explicit content without their consent. This manipulation utilized GAN-based software, such as FaceApp, trained on extensive datasets of the celebrities' images. Detection efforts focused on identifying artifacts such as unnatural facial movements and lighting inconsistencies in the manipulated content.

D. Comparative Analysis

The Pelosi deepfake aimed to manipulate political perceptions, highlighting how generative AI can alter political narratives. The BBC experiment served as a controlled illustration of deepfake technology's capabilities and risks, whereas deepfake pornography cases exemplify the exploitation and privacy violations enabled by these technologies.

Detection of the Pelosi deepfake utilized traditional forensic techniques, while the BBC's experiment underscored the effectiveness of advanced AI-based detection tools. In contrast, deepfake pornography detection relied on sophisticated image analysis to uncover visual inconsistencies.

The impact of manipulated videos on public trust and perception is profound, emphasizing the need for robust detection mechanisms. The use of generative AI for non-consensual explicit content raises significant ethical and legal concerns, underscoring the necessity for comprehensive regulatory frameworks. As AI technology advances, continuous innovation in forensic methods and detection tools remains essential.

These case studies illustrate the transformative power of generative AI while highlighting the pressing need for advanced detection strategies and regulatory measures to address the associated risks and ethical issues.

VIII. Discussion

The empirical case studies and technical evaluations presented in this paper reveal critical gaps in the current capabilities of forensic tools to detect AI-manipulated video content. Specifically, deepfake detection techniques, such as artifact detection and deep learning-based approaches, exhibit high rates of false negatives when faced with increasingly sophisticated manipulation methods, including real-time facial reenactment and audio-visual synchronization. These findings highlight the limitations of existing detection technologies, which are often unable to keep pace with the rapid advancements in generative AI.

The case studies, including high-profile examples such as the manipulated video of U.S. Speaker Nancy Pelosi, illustrate the far-reaching implications of deepfake technology in undermining public trust in video evidence. In this case, traditional forensic techniques were inadequate in identifying the deepfake, necessitating the development of more advanced detection mechanisms. The "BBC Deepfake Experiment" further emphasizes the challenge by demonstrating how AI-generated content can convincingly mimic reality, even in controlled environments. These real-world examples underscore the growing need for improved detection capabilities to safeguard the integrity of video evidence in legal and investigative contexts.

Moreover, a critical review of the literature reveals a significant gap between the current state of deepfake detection technologies and the standards required for the admissibility of digital evidence in courtrooms. Current legal frameworks, which traditionally rely on the presumption of video authenticity, are increasingly inadequate in dealing with AI-generated content. As the sophistication of generative AI grows, so too does the risk of wrongful convictions or acquittals due to undetectable manipulations in video evidence.

In light of these findings, this paper argues that a multidisciplinary approach is essential to address the challenges posed by AI-manipulated media. Advancements in forensic technologies must be accompanied by legal reform to establish robust evidentiary standards capable of adapting to the evolving landscape of digital media. Without this integrated approach, both legal systems and forensic methods will continue to lag behind the capabilities of generative AI, compromising the credibility of video evidence.

IX. Future Directions and Recommendations

A. Predictions for Generative AI and Associated Risks

Generative AI is evolving rapidly, driven by advancements in algorithms, computational power, and data. Key trends and risks include:

- Increased Realism and Accessibility: Future models will generate highly realistic content, making it challenging to distinguish between genuine and synthetic media. The widespread availability of these tools may also facilitate malicious uses.
- Integration with Emerging Technologies: Generative AI is likely to integrate with AR, VR, and IoT, creating immersive and potentially manipulative content, which could blur the lines between reality and virtual media.

- Ethical and Societal Challenges: Enhanced capabilities will escalate misinformation and disinformation risks, undermining public trust and raising privacy concerns due to the misuse of personal data.

B. *Recommendations for Mitigating Risks*

- Detection and Verification: Invest in advanced detection technologies and establish industry-wide standards to ensure reliable identification of manipulated content.
- Policy and Regulation: Develop comprehensive legislation to address the creation and misuse of manipulated videos, incorporating strict privacy protections and penalties for malicious use.
- Transparency and Accountability: Require disclosure of synthetic content and implement mechanisms to hold creators accountable for malicious activities.
- Public and Professional Education: Launch awareness campaigns to educate the public on AI-generated content and provide specialized training for legal and forensic professionals.

C. *Enhancing Video Evidence Reliability*

- **Digital Watermarking and Blockchain:** Use digital watermarks for traceability and blockchain for immutable records to ensure video authenticity and integrity.
- **AI-Augmented Verification:** Develop AI tools for verifying video content against databases and integrate cross-platform systems for comprehensive verification.
- **Collaboration and Standards:** Foster collaboration among technology companies, regulators, and researchers to develop global standards for video evidence verification and manipulation detection.

These measures are crucial for addressing the challenges of generative AI and ensuring the integrity of video evidence in an evolving technological landscape.

X. CONCLUSION

This paper has explored the profound implications of generative AI, particularly deepfake technologies, on the reliability of video evidence in legal and investigative contexts. Through an analysis of existing detection methods and a review of high-profile case studies, it is clear that current forensic tools are insufficient to meet the challenges posed by sophisticated AI-generated content. The findings underscore the urgent need for more advanced detection techniques, alongside a reevaluation of legal standards governing the admissibility of digital evidence.

The results of this study reveal not only the technological gaps but also the legal and ethical challenges that must be addressed to preserve the integrity of video evidence. As generative AI technologies continue to evolve, their potential for misuse will likely increase, raising critical concerns for both law enforcement and the judicial system. It is therefore imperative that future research focuses on developing robust detection mechanisms that can keep pace with AI-driven media manipulation.

In conclusion, this paper advocates for a comprehensive approach that combines technological innovation, legal reform, and ethical guidelines to mitigate the risks posed by generative AI. Collaboration among technologists, legal experts, and policymakers will be crucial in ensuring that the integrity of video evidence is maintained in an increasingly digital world. Future research should aim to bridge the gap between forensic capabilities and legal standards, ensuring that video evidence remains a trusted component of the judicial process.

REFERENCES

[1] C. Law, "Journal of Digital Forensics, Security and Law," vol. 6, no. 3, 2011

[2] R. Chesney and D. Citron, "Deepfakes and the new disinformation war: The coming age of post-truth geopolitics," Foreign Affairs, vol. 98, p. 147, 2019.

[3] A. Godulla, C. P. Hoffmann, and D. Seibert, "Dealing with deepfakes—an interdisciplinary examination of the state of research and implications for communication studies," SCM Studies in Communication and Media, vol. 10, no. 1, pp. 72-96, 2021.

[4] I. Goodfellow et al., "Generative adversarial nets," in Advances in Neural Information Processing Systems, vol. 27, 2014.

[5] P. Korshunov and S. Marcel, "Deepfakes: A new threat to face recognition? Assessment and detection," arXiv preprint arXiv:1812.08685, 2018.

[6] L. T. Phan et al., "Importation and human-to-human transmission of a novel coronavirus in Vietnam," New England Journal of Medicine, vol. 382, no. 9, pp. 872-874, 2020.

[7] N. Aldausari, "Audio-to-Video Synthesis using Generative Adversarial Network," Ph.D. dissertation, University of New South Wales, Sydney, Australia, 2024.

[8] V. Singh, A. Dixit, S. Pandey, B. V. Kumar, V. Pachouri, and M. Sahu, "Role of Artificial Intelligence in Criminal Investigation," in 2023 International Conference on Quantum Technologies, Communications, Computing, Hardware and Embedded Systems Security (iQ-CCHESS), Sep. 2023, pp. 1-4. IEEE.

[9] S. Lyu, "Deepfake detection," in Multimedia Forensics, pp. 313-331. Singapore: Springer Singapore, 2022.

[10] A. Agarwal, A. Agarwal, S. Sinha, M. Vatsa, and R. Singh, "MD-CSDNetwork: Multi-domain cross stitched network for deepfake detection," in 2021 16th IEEE International Conference on Automatic Face and Gesture Recognition (FG 2021), Dec. 2021, pp. 1-8. IEEE.

[11] W. Lu, L. Liu, B. Zhang, J. Luo, X. Zhao, Y. Zhou, and J. Huang, "Detection of deepfake videos using long-distance attention," IEEE Transactions on Neural Networks and Learning Systems, 2023.

[12] A. Sufian, "AI-Generated Videos and Deepfakes: A Technical Primer," Authorea Preprints, 2024.

[13] I. Chaudhuri and B. D. Hemmige, "DEEP FAKE: An emerging threat in the digital media manipulation," Sampreshan, vol. 17, no. 2, pp. 833-861, ISSN: 2347-2979, UGC CARE Group 1, 2024.

[14] T. Wang, H. Cheng, K. P. Chow, and L. Nie, "Deep convolutional pooling transformer for deepfake detection," ACM Transactions on Multimedia Computing, Communications, and Applications, vol. 19, no. 6, pp. 1-20, 2023.

[15] S. D. Nelson, J. W. Simek, and M. Maschke, "Detecting deepfakes," Law Practice, vol. 46, p. 42, 2020.

[16] C. Vaccari and A. Chadwick, "Deepfakes and disinformation: Exploring the impact of synthetic political video on deception, uncertainty, and trust in news," Social Media+ Society, vol. 6, no. 1, p. 2056305120903408, 2020.

Adaptive Traffic Control Framework for Urban Intersections

K.K.S.S.Wijayarathna
Faculty of Science
NSBM Green University
Sri Lanka.
kksswijayarathna@students.nsbm.ac.lk

H.K.I.S.Lakmal
Faculty of Science
NSBM Green University
Sri Lanka.
isuru.l@nsbm.ac.lk

***Abstract*— Traffic congestion in Colombo, particularly at heavily used intersections, has been a persistent challenge for decades, largely due to the increasing number of vehicles and the limitations of current traffic management systems. The existing fixed-time traffic signals and manual interventions fail to adapt to the complex and evolving traffic patterns, leading to extended vehicle queues, prolonged waiting times, and significant commuter dissatisfaction. This research addresses these issues by proposing an adaptive traffic signal control framework for a major intersection in Colombo. The system utilizes real-time data collection from a CCTV camera installed near the intersection to monitor traffic patterns continuously. An LSTM network is employed to predict traffic volumes based on the identified patterns, while a Deep Deterministic Policy Gradient (DDPG)-based reinforcement learning algorithm is used to optimize traffic light timings. The proposed solution aims to maximize vehicle throughput, reduce waiting times, and minimize vehicle queues, while also being easily integrated into existing infrastructure. The framework is designed to be both economically feasible, requiring reduced processing power and offering a robust solution for urban traffic management.**

Keywords - Adaptive traffic control, Reinforcement Learning (RL), Deep Neural Networks

I. INTRODUCTION

Traffic congestion is an escalating problem in urban areas worldwide, adversely impacting daily life, economic productivity, and environmental sustainability. As urban populations grow and vehicle numbers increase, traffic bottlenecks become more frequent and severe. Traditional traffic management systems, which rely on static, fixed-time traffic signals, are often inadequate for addressing the dynamic and complex traffic patterns observed in modern urban environments.

In Colombo, Sri Lanka's capital, traffic congestion has long been a persistent issue, particularly at key urban intersections that are vital nodes in the city's transportation network. These intersections, where vehicles from multiple directions converge, are often the most congested points, leading to complex traffic scenarios and longer waiting times. The current traffic control system in Colombo, which combines fixed-time signal controllers with manual interventions, is inadequate for handling the growing traffic volumes and complex flows, necessitating frequent manual adjustments during peak hours. This inefficiency underscores the need for an adaptive traffic control system that can intelligently adjust signal timings in real-time based on actual traffic conditions.

Existing approaches in Colombo have not sufficiently addressed the need for real-time adaptability in traffic management. Although CCTV cameras monitor some intersections, these systems are primarily for recording footage rather than active traffic management. Moreover, the high costs associated with installing multiple cameras to monitor a single intersection make widespread deployment impractical, especially in a developing country like Sri Lanka during economic challenges. This research seeks to fill this gap by developing an adaptive traffic signal control system specifically tailored to a major intersection in Colombo, proposing a cost-effective and scalable framework for implementation across Sri Lanka.

The proposed solution integrates real-time data collection, advanced neural networks, and reinforcement learning algorithms to dynamically optimize traffic signal timings. A CCTV camera installed at the intersection continuously monitors traffic patterns, with the data feeding into a Long Short- Term Memory (LSTM) network designed to predict traffic volumes based on both historical and real-time inputs. These predictions are used by a Deep Deterministic Policy Gradient (DDPG)-based reinforcement learning model to adjust traffic light sequences, maximizing vehicle throughput, reducing waiting times, and minimizing queues. A key feature of this approach is the ability to test the framework in a SUMO simulation environment before real-world implementation which will be able to cut-off unexpected cost that have to be bared with prototypes development.

The novelty of this approach lies in its adaptability to real-time conditions and its economic feasibility for a developing country like Sri Lanka. The system is designed to require minimal infrastructure and processing power, making it cost-effective and easily integral into existing traffic management frameworks. By addressing

the limitations of current traffic management strategies and offering a scalable, cost- effective solution, this research has the potential to significantly

improve traffic flow and reduce congestion in Colombo's urban environment.

This study aims to establish a robust framework for adaptive traffic signal control that can be implemented in other intersections across Colombo and similar urban areas facing congestion challenges. Expected outcomes include enhanced traffic management efficiency, reduced environmental impact, and improved commuter satisfaction, contributing to sustainable urban development.

II. Related Work

Traffic signal control has undergone significant evolution over the years, driven by the growing need to manage increasing congestion and optimize traffic flow in urban environments. This section reviews various approaches to traffic signal control, focusing on static systems, adaptive methods, and the application of reinforcement learning (RL) techniques.

A. Static Traffic Signal Control

Static traffic signal control remains prevalent in many regions, including Colombo, Sri Lanka. In these systems, traffic signal parameters—such as phase durations and cycle times—are predetermined based on historical traffic data. While effective under stable traffic conditions, but struggle with dynamic and unpredictable scenarios, such as adverse weather, high traffic conditions. These limitations often lead to inefficiencies, as the fixed timings cannot adapt to changing traffic volumes and conditions.

B. Adaptive Traffic Signal Control

To address the limitations of static traffic systems, adaptive traffic signal control has emerged as a promising solution. Various adaptive methods have been proposed:

Genetic Algorithms (GAs): GAs optimize traffic signal settings by treating phases and timings as chromosomes, aiming to minimize queue lengths and travel times through a fitness function [3][4]. However, GAs can be computationally intensive and slow to converge.

Fuzzy Logic: This method uses linguistic variables and fuzzy inference rules to model uncertainty and determine signal timings based on imprecise data [5][6]. While useful for handling uncertainty, fuzzy logic systems can become complex and less adaptable to highly dynamic traffic conditions.

Reinforcement Learning (RL): Q-learning, a popular RL approach, learns optimal signal timings based on vehicle queues and actions that adjust green times [7]. Despite its advantages, Q-learning is limited by its discrete action space, restricting its ability to manage continuous traffic variations effectively.

Swarm Intelligence: Techniques like Ant Colony Optimization (ACO) mimic social insect behaviour to solve combinatorial problems in traffic control [1]. While effective in exploring complex search spaces, ACO may struggle with real-time applications due to high computational demands.

C. Advancements in Reinforcement Learning for Adaptive Traffic Control

Recent advancements in adaptive traffic control, particularly through reinforcement learning (RL), have significantly improved vehicle flow optimization in

complex urban environments. Traditional traffic signal control systems, such as Webster's Formula and SCATS, relied on static models and struggled with unpredictable traffic patterns. RL provides a dynamic, data-driven alternative, allowing systems to learn optimal strategies by interacting with their environment in real time. The foundational framework of RL is rooted in the Markov Decision Process (MDP), where agents learn by taking actions based on observed states and maximizing cumulative rewards. Q-learning is a popular RL technique for estimating long-term value, and the integration of Deep Reinforcement Learning (DRL) has further enhanced its capabilities. DRL enables traffic control systems to handle complex environments and process high-dimensional data, such as real-time sensor feeds, without manual feature extraction.

RL-based methods have shown great potential in improving adaptive traffic control by dynamically adjusting to real-time traffic conditions. [11] demonstrated how Q-learning could outperform traditional methods by learning from real-time traffic patterns at isolated intersections, with the potential for expansion to networked systems. Similarly, [12] surveyed the use of multi-agent reinforcement learning (MARL), where intersections act as independent agents that cooperate to optimize traffic across urban networks. This decentralized approach is crucial for managing interconnected urban intersections.

DRL's ability to process complex data, such as visual inputs from traffic cameras, allows adaptive systems to adjust signal timings based on traffic patterns. Real- world implementations, such as the Surtrac system in Pittsburgh, have shown success in reducing wait times and emissions. Additionally, [11] reported success in applying RL to an isolated traffic signal in Toronto, with real-time adaptation resulting in improved traffic flow. [12] emphasized the importance of sophisticated simulation environments in validating RL models before deployment, ensuring scalability and adaptability across various traffic conditions.

D. DDPG Reinforcement Learning Approach

The Deep Deterministic Policy Gradient (DDPG) algorithm offers significant advantages for real-time traffic signal control, as highlighted in Wu's research. One of the key strengths of DDPG is its capability to handle continuous action spaces, which is essential for making fine-tuned decisions regarding traffic signal timings (e.g., adjusting green, yellow, and red light durations). In contrast, traditional reinforcement learning (RL) algorithms, such as Q- learning and Deep Q-Networks (DQN), are limited to discrete action spaces, which can hinder their effectiveness in complex traffic scenarios.

DDPG's actor-critic architecture allows it to operate in high-dimensional state and action spaces,

effectively processing multi-dimensional inputs like vehicle counts, waiting times, and road occupancy. This enables the development of sophisticated policies that can adapt to various traffic patterns, unlike simpler RL algorithms that may struggle in such environments.

Wu's research also indicates that DDPG achieves faster convergence and more stable performance compared to DQN and Normalized Advantage Function (NAF) algorithms in simulations. The use of experience replay in DDPG further enhances learning efficiency, allowing the agent to leverage past experiences for improved policy robustness, which is crucial for effective traffic management. DDPG's ability to manage continuous action spaces and multi-dimensional state variables positions it as an ideal candidate for real-time traffic signal control, demonstrating its potential to mitigate congestion and enhance overall traffic flow.

E. Camera Based Traffic Detection Systems

The increasing complexity of urban traffic necessitates effective solutions for real-time vehicle detection and identification. Camera-based systems, particularly those using advanced deep learning models, have proven reliable for traffic monitoring.[10] emphasize the vital role of Closed Circuit Television (CCTV) cameras in Intelligent Transportation Systems (ITS), enhancing incident detection and vehicle monitoring while improving traffic flow management.

The YOLOv8 model stands out for its accuracy, achieving a training accuracy of 77% and a testing accuracy of 96% on a diverse vehicle dataset. Its robust architecture effectively handles challenges such as occlusion and overlapping vehicles. Combining YOLOv8 with CCTV technology shows great potential for real-time vehicle detection, as its rapid image processing capabilities facilitate timely responses to traffic conditions. Data augmentation techniques further improve detection accuracy, particularly for small objects. In this study, utilizing a CCTV camera positioned at a distance to monitor intersection vehicle activity effectively leverages YOLOv8's strengths, offering a reliable solution to real-time vehicle detection challenges and ensuring accurate urban traffic management.

F. Traffic Prediction systems

Long Short-Term Memory (LSTM) networks have proven to be a highly effective tool for traffic flow prediction, as highlighted by [9]. Their research demonstrates that LSTMs excel at modeling complex

temporal dependencies in traffic data, resulting in superior forecasting accuracy compared to traditional methods like random walk and support vector regression. The unique architecture of LSTMs, which incorporates memory cells and gating mechanisms, allows them to capture intricate traffic patterns over time.

Additionally, LSTM implementation can lead to significant cost savings by reducing the need for multiple camera installations at intersections. Instead, a single camera, combined with date and time data, can be used to estimate traffic volumes on other roads. This approach not only lowers installation and maintenance costs but also simplifies data collection, making it an efficient solution for traffic management. The successful application of LSTM networks in traffic flow prediction underscores their reliability and cost- effectiveness, enabling traffic authorities to optimize resource allocation and enhance overall traffic flow efficiency.

III. PROPOSED SYSTEM DESIGN

A. Proposed Framework

The proposed adaptive traffic signal control system is designed to mitigate traffic congestion at urban intersections by utilizing advanced machine learning and reinforcement learning techniques. The system is composed of three primary components: a YOLO model for vision-based real-time vehicle detection, a Long Short-Term Memory (LSTM) network for traffic flow prediction, and a Deep Deterministic Policy Gradient (DDPG)-based reinforcement learning algorithm for optimizing traffic signal timings.

B. Data Flow in the Proposed Framework

The data flow in the proposed system initiates with the acquisition of a real-time video feed from a single CCTV camera installed at the intersection. This camera is placed to monitor traffic conditions on one of the roads leading to the intersection. The captured video feed is processed in real time by the YOLO model. YOLO model extracts essential information. This processed data, along with supplementary contextual information such as the date and time, is subsequently input into the LSTM network.

The LSTM network is tasked with predicting traffic flow on the other three roads that converge at the intersection. By analyzing the processed data from the YOLO model, the LSTM network generates accurate traffic predictions, leveraging historical traffic patterns and real-time data. This prediction capability significantly reduces the need for multiple cameras and the associated processing power, as it enables the system to predict traffic conditions on all roads of the intersection using data from a single camera.

Following the traffic predictions, the output data is transferred to the DDPG-based reinforcement learning (RL) model. The RL model employs these predictions to determine the optimal traffic light sequence for the intersection. The primary objectives are to maximize vehicle throughput, minimize waiting times, and reduce vehicle queues by

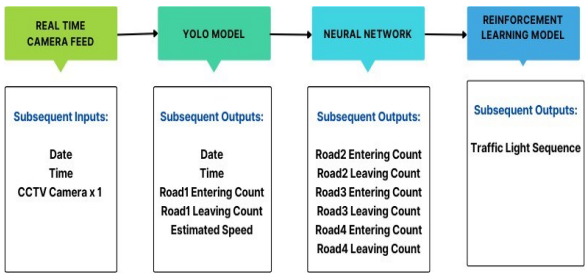


Fig. 1. Flow of proposed framework

dynamically adjusting the signal timings based on real-time traffic conditions. This integration of the YOLO model, LSTM network, and DDPG-based RL model into a unified framework creates a cost-effective and adaptive traffic management system capable of responding to the ever- changing demands of urban traffic.

IV. METHODOLOGY

A. Traffic Perception Survey

Prior to selecting the Borella intersection for this study, a comprehensive traffic perception survey was conducted to ensure the site accurately reflected typical traffic conditions and challenges. The survey, distributed online to a random sample of intersection users, collected preliminary data on traffic patterns, peak hours, and common congestion issues. Additionally, data from the Borella Police Station provided critical insights into traffic volume and complexity at the intersection. The combined findings confirmed Borella as a high-demand area with intricate traffic flows, making it an optimal focus for this research. Fig 2 illustrates key concerns and traffic patterns from the survey.

B. Data Collection

The data collection for this study was carried out in two distinct phases to ensure the accuracy and reliability of the dataset. The initial phase involved the use of CCTV footage provided by the Colombo City Police's CCTV division. However, inconsistencies in the quality and coverage of the footage made it necessary to conduct a second phase of data collection to obtain a more reliable dataset.

In the second phase, a V380 Pro WiFi outdoor camera was strategically installed at a construction site near the Borella intersection for continuous monitoring over a three-week period. Prior to installation, necessary permissions were obtained from the building owners and both the Borella and Colombo North Police Stations to ensure compliance with

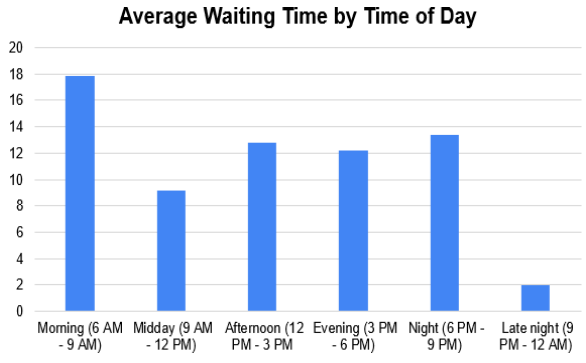


Fig. 2. Insights from traffic demand analysis

local regulations. This study adopted a novel approach by using a single- camera setup, chosen for its cost-effectiveness and feasibility. The V380 Pro camera, equipped with a wide-angle dual-lens, provided comprehensive coverage of the entire intersection from a distance, reducing the need for multiple cameras. The camera was fitted with a 128GB

Fig. 3. Camera installed at the junction

micro SD card, offering ample storage capacity for extended recording. Additionally, the camera's real-time monitoring capabilities allowed remote access to check its status and manage SD card usage via a dedicated mobile application, ensuring continuous oversight throughout the data collection process.

The V380 Pro camera's night vision functionality also provided clear footage during nighttime, which was a significant improvement over the police surveillance cameras used in the first phase. The camera setup required only a power source and a SIM card with sufficient data capacity to maintain connectivity, making it a cost-effective and practical solution for capturing traffic data at the intersection. The recorded video files were saved in AVI format and later converted to MP4 for further analysis. VLC Media Player was used for efficient transcoding, ensuring the video files were in a more accessible format for processing and analysis.

C. Vehicle Detection model

The collected CCTV data will be employed to train a vehicle detection and counting model based on YOLOv8l. This proposed adaptive traffic signal control system integrates the YOLO (You Only Look Once) model for real-time vehicle detection and tracking, which is essential for effectively managing traffic flow at urban intersections. Utilizing the ultralytics/yolo framework, the YOLOv8m model is optimized for high-speed processing, enabling real-time classification of detected objects, including cars, buses, and motorcycles.

To enhance detection capabilities, the YOLO model is combined with the Deep SORT (Simple Online and Realtime Tracking with a Deep Association Metric) algorithm. This integration allows the system not only to detect vehicles but also to track them across multiple frames, assigning a unique identifier to each vehicle that is maintained until the vehicle exits the scene.

Additionally, the system incorporates vehicle counting and speed estimation by defining reference lines at the intersection. These lines are used to count vehicles as they enter and exit different roads. Vehicle speed is estimated by analyzing movement across frames in relation to these reference lines. The system also captures and utilizes contextual data, including date, time, and estimated vehicle speed. To optimize the YOLO model for the Borella intersection, video data covering both daytime and nighttime traffic, including peak and off-peak hours, as well as challenging conditions such as adverse weather, will

be extracted and annotated. A custom labelling system was developed to differentiate between vehicle appearances during day and night. This approach enhances the model's capability to detect and track vehicles under varying lighting conditions.

The annotated dataset will be used to train the YOLO model, optimizing its performance for detecting and tracking vehicles across diverse lighting and traffic conditions. A key feature of this YOLO model is its customization for the Borella intersection, ensuring efficient vehicle counting and tracking during both daytime and nighttime. This specificity enhances the model's robustness and accuracy in various lighting scenarios, which is critical for maintaining consistent performance throughout the day and night.

D. Traffic Demand Analysis

Following the collection of traffic data via camera footage, a comprehensive traffic demand analysis was performed to evaluate patterns at the Borella intersection. The YOLO model was used to accurately detect and classify vehicles, enabling a detailed assessment of traffic volumes, peak periods, and daily flow variations. This analysis identified high and low traffic periods, providing crucial insights into congestion levels throughout the day.

E. Traffic Flow Prediction model

The proposed system's traffic volume prediction module leverages a Long Short-Term Memory (LSTM) network, a specialized type of recurrent neural network (RNN) optimized for time-series data analysis. This LSTM model predicts traffic volumes at an intersection by utilizing real-time data from one road and drawing on patterns learned from historical traffic data. This approach minimizes the need for multiple cameras, thus reducing monitoring costs. Developed using TensorFlow and Keras, the model's architecture is fine- tuned through Keras Tuner's RandomSearch algorithm for optimal hyperparameter settings.

The input data, derived from the YOLO model outputs (date, time, and traffic volumes), is processed through six traffic-related metrics and standardized using a StandardScaler to enhance training efficiency. These inputs are then reshaped to match the LSTM's required format. Extensive hyperparameter tuning minimizes the mean absolute error (MAE) on a validation set. To prevent overfitting and improve performance, strategies like early stopping and model checkpointing are employed during training. Once trained, the network predicts traffic volumes on other roads at the intersection based on real-time input and learned historical patterns.

F. Traffic Control Algorithm

The core of the decision-making process in the proposed system is powered by the Deep Deterministic Policy Gradient (DDPG) algorithm, a model-free, off- policy reinforcement learning approach well-suited for continuous action spaces. Unlike Q-learning, which is designed for discrete actions, DDPG excels in environments where precise control over actions is required, such as determining optimal traffic light durations. DDPG integrates elements from Deep Q- Learning (DQL) and Policy Gradient methods, utilizing two neural networks: an actor network, which outputs a deterministic action (e.g., the duration of a green light), and a critic network, which evaluates this action by predicting the expected future rewards. The DDPG algorithm follows key steps:

- Experience Replay: The agent stores experiences (state, action, reward, next state) in a replay buffer.
- Actor-Critic Framework: The actor network outputs continuous actions based on the current state, while the critic network evaluates these actions by predicting the Q- value, which reflects the expected cumulative reward, such as reduced congestion and waiting times.
- Target Networks: To stabilize training, target networks are employed for both actor and critic networks. These target networks are slowly updated to follow the main networks, ensuring gradual and stable updates to the Q- value.
- Policy Update: The actor network's policy is updated by optimizing the expected cumulative reward, encouraging actions that maximize future rewards. The critic network minimizes the loss between the predicted and target Q-values, calculated using the Bellman equation:

  $$Q(st,at) = rt+1+\gamma \cdot Q'(st+1,\mu'(st+1))$$

  where Q(st,at) is the estimated Q-value for the current state-action pair, rt+1 is the reward and $Q'(st+1, \mu'(st+1))$ is the Q-value predicted by the target networks.

- Exploration vs. Exploitation: To ensure that the agent explores a wide range of actions during training, noise is added to the actions generated by the actor network.

The proposed adaptive traffic signal control framework leverages the Deep Deterministic Policy Gradient (DDPG) algorithm to optimize traffic light sequences at the intersection. The primary objective is to enhance traffic flow, minimize vehicle waiting times, and reduce queue lengths.

- State: In the DDPG model, the state is a multidimensional vector representing real- time traffic conditions at the intersection. It includes the number of vehicles on each incoming lane, waiting time and queue lengths.
- Action: The action involves the agent determining the traffic light time sequences, specifically the duration for which each light remains green, yellow, or red. Given that these durations are continuous variables.
- Reward: The reward function is crafted to balance various aspects of traffic management, encouraging the agent to optimize the overall traffic flow. The reward is calculated as:

  R=w1·(Number of vehicles passing through t he intersection)−w2·(Average waiting time)−w3·(Average queue length)

  where w1, w2, and w3 are the weights assigned to each component. The goal is to maximize

vehicle throughput while minimizing waiting times and queue lengths.

The DDPG model's output is a priority-based green light time sequence, dynamically adjusted based on real-time traffic volumes. Roads with higher traffic volumes receive extended green light durations, ensuring that congestion is effectively mitigated. The final output is an optimized sequence of green light timings for each road at the Borella intersection, designed to maximize vehicle throughput while

minimizing delays and congestion. This approach harnesses the power of reinforcement learning in continuous action spaces to provide an adaptive and sophisticated traffic management solution.

G. Simulation Model

The Simulation of Urban Mobility (SUMO) environment played a crucial role in developing and testing the proposed adaptive traffic signal control system. SUMO served two primary functions: it acted as a platform for training the Deep Deterministic Policy Gradient (DDPG) reinforcement learning algorithm and as a testing ground for evaluating the RL model using real-world traffic data. In reinforcement learning, an environment is essential for the agent to learn from the consequences of its actions. SUMO offers a controlled, risk-free setting to simulate various traffic scenarios, enabling the RL model to experiment with different traffic light sequences and assess their effectiveness prior to real-world deployment.

The system's performance will be compared against traditional static pre-timed traffic light controllers, focusing on metrics such as vehicle throughput, average waiting time, and queue lengths. SUMO's capability to accurately simulate and measure these metrics, especially queue lengths—which are challenging to capture via cameras due to extensive queuing distances at Borella—provides valuable insights. This simulation-based approach ensures that the RL model is rigorously tested, refined, and prepared to manage real-world traffic conditions effectively, thereby minimizing the risks associated with direct implementation.

Fig. 4. RL model training process

V. RESULTS

Initial progress has been made in several key areas of the project. The data collection phase was successfully completed, with 24-hour monitoring using a Wi-Fi CCTV camera installed at the Borella intersection over three continuous weeks, providing real-world data records. The footage obtained is indicated by Fig.6.

The YOLO model was successfully developed and tested, demonstrating its ability to detect vehicles, count them, and record timestamps from the captured footage. Further fine-tuning was conducted using footage from the camera installed at the Borella intersection to account for its specific conditions. After annotating over 300-500 frames, the model was retrained to handle varying traffic volumes, weather conditions such as rain, and different lighting scenarios, including daylight and night conditions, until the desired performance was achieved.

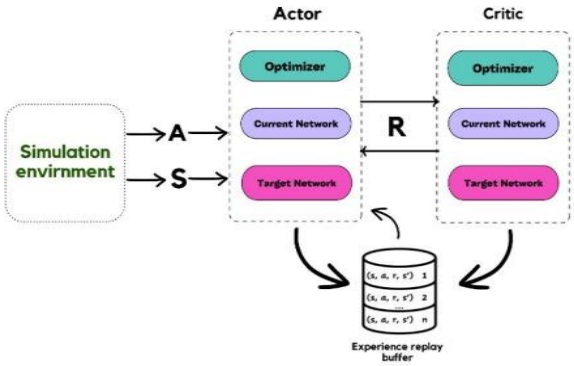


Fig. 6. Frame Obtained from the CCTV video

Figure 7 presents the results of the traffic demand analysis, highlighting the identified traffic patterns, peak hours, and flow variations throughout the day.

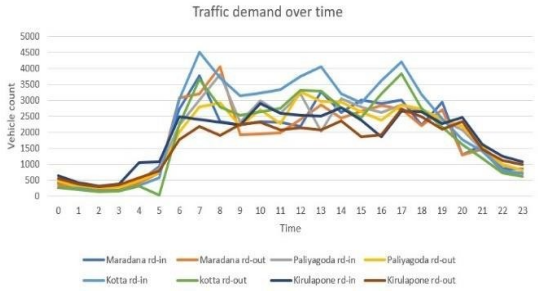


Fig. 7. Traffic Demand Analysis

Fig. 8. Processed frame from YOLO model

Fig. 9. Annotated Frame

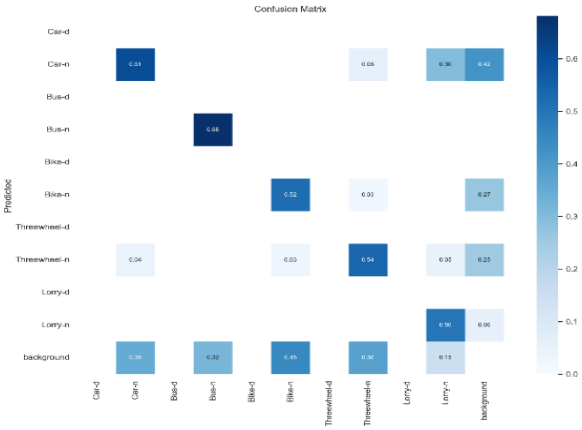


Fig. 10. Confusion Matrix of YOLO mod

An LSTM network was initially developed using a dataset sourced from Kaggle until the completion of real-world data collection. While the model showed promising initial results, it was subsequently fine-tuned with extracted data from real-world camera footage specific to the Borella junction to enhance its accuracy and applicability. The results obtained from this refinement are illustrated in Fig 10.

The SUMO network representing the Borella intersection has been successfully constructed using OpenStreetMap data. A Deep Deterministic Policy Gradient (DDPG) based reinforcement learning algorithm has been developed; however, the model remains in the exploration phase. Further fine-tuning is necessary to optimize the algorithm for effectively exploiting actions to achieve the desired outcomes.

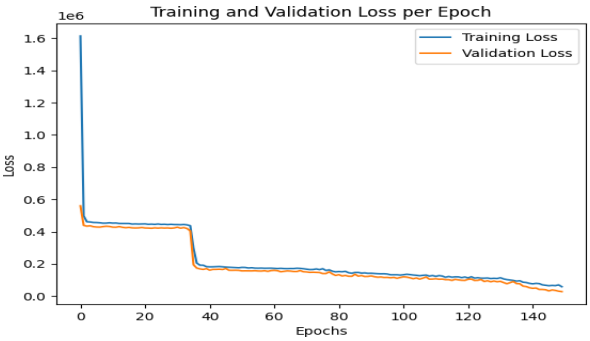


Fig. 11. Performance of LSTM network

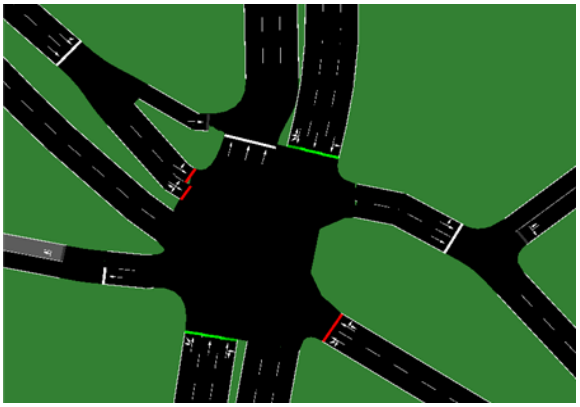


Fig. 12. SUMO network of Borella junction

VI. Conclusion

This study proposes an innovative adaptive traffic signal framework for urban intersections and conducts a quantitative evaluation to assess its feasibility for real-world implementation. Future research should focus on real-world deployment of the proposed framework, improving its capability to handle complex scenarios, and exploring its scalability across multiple intersections for wider application.

Acknowledgment

I would like to express my sincere gratitude to Mr. Janaka Rathnayaka for his invaluable support in providing the building space and power facilities necessary for the successful installation of the camera used in this research. His assistance was instrumental in enabling the data collection process, and I am deeply appreciative of his contributions to the successful execution of this study.

References

[1] J. Withanawasam and A. Karunananda, “Multi-agent based road traffic control optimization,” in 2017 IEEE 20th International Conference on Intelligent Transportation Systems (ITSC), 2017.

[2] M. Sutharsan, S. Rajakaruna, S. Y. Jayaweera, J. A. C. M. Jayaweera, and S. Thayaparan, "Vision-based adaptive traffic light controller for single intersection," in 2020 5th International Conference on Information Technology Research (ICITR), 2020.

[3] K. Shingate, K. Jagdale, Y. Dias, and Fr. c Rodrigues Institute of Technology, "Adaptive traffic control system using reinforcement learning," Int. J. Eng. Res. Technol. (Ahmedabad), vol. V9, no. 02, 2020

[4] H. Wu, "Control method of traffic signal lights based on DDPG reinforcement learning," J. Phys. Conf. Ser., vol. 1646, no. 1, p. 012077, 2020.

[5] T. Mahmood, M. E. M. Ali, A. Durdu, and Electrical and Electronic Engineering, Selçuk University, Konya, Turkey, "A two stage fuzzy logic adaptive traffic signal control for an isolated intersection based on real data using SUMO simulator," International Journal of Trend in Scientific Research and Development, vol. 3, no. 4, pp. 656–659, 2019.

[6] W. A. C. J. K. Chandrasekara, R. M. K. T. Rathnayaka, and L.L. G. Chathuranga, "A real-time density-based traffic signal control system," in 2020 5th International Conference on Information Technology Research (ICITR), 2020.

[7] Y.-J. Lee and O. Min, "Long short-term memory recurrent neural network for urban traffic prediction: A case study of Seoul," in 2018 21st International Conference on Intelligent Transportation Systems (ITSC), 2018.

[8] R. T. Wakkumbura, B. Hettige, and A. Edirisuriya, "Real-time traffic controlling system using Multi-Agent technology," J. Eur. Syst. Autom., vol. 54, no. 4, pp. 633–640, 2021.

[9] Shao, H., & Soong, B.-H. Traffic flow prediction with Long Short-Term Memory Networks (LSTMs). Interdisciplinary Graduate School, Nanyang Technological University.

[10] Hati Telaumbanua, A. P., Larosa, T. P., Pratama, P. D., Fauza, R. H., & Husein, A. M. (2024). Vehicle Detection and Identification Using Computer Vision Technology with the Utilization of the YOLOv8 Deep Learning Method. Universitas Prima Indonesia, Fakultas Sains dan Teknologi, Kota Medan.

[11] Abdulhai, B., Pringle, R., & Karakoulas, G. J. (2003). Reinforcement Learning for True Adaptive Traffic Signal Control. Journal of Transportation Engineering, 129(3), 278-285.

[12] Wei, H., Zheng, G., Gayah, V., & Li, Z. (2021). Recent Advances in Reinforcement Learning for Traffic Signal Control: A Survey of Models and Evaluation.

Machine Learning Techniques for Predicting Brain Stroke Risk: Addressing Data Imbalance

L B H C Pathmakumara
University of Plymouth
United Kingdom
hchandeepap@gmail.com

R W K T Rajapaksha
Faculty of Computing
NSBM Green University
Sri Lanka
kavishka.r@nsbm.ac.lk

Abstract—**Stroke represents a significant global health concern, markedly influencing morbidity and mortality rates across the world. This research investigates the utilization of various machine learning algorithms to forecast stroke risk, with a specific emphasis on tackling the widespread issue of class imbalance found in stroke prediction datasets. We conduct a thorough assessment of the performance of several models, including Random Forest, K-Nearest Neighbors (K-NN), Naive Bayes, Decision Trees, Support Vector Machines (SVM), and Logistic Regression, analyzing both the original imbalanced dataset and a balanced dataset created using the Synthetic Minority Over-sampling Technique (SMOTE). Our results indicate that although the initial accuracy of the models on the imbalanced dataset was substantial, the implementation of SMOTE provided a more accurate evaluation of model efficacy, with Random Forest attaining the highest accuracy at 92%. This study highlights the critical role of class balancing methods such as SMOTE in improving the predictive capabilities of machine learning models within medical applications. The findings of this research carry considerable implications, as enhanced stroke prediction models can facilitate better patient outcomes and optimize healthcare resource allocation through timely interventions.**

Keywords— stroke prediction, machine learning, class imbalance, SMOTE, healthcare optimization

I. INTRODUCTION

A stroke is a serious disease with a high chance of death that arises when the brain's blood flow is partially interrupted. It ranks as the second leading cause of mortality and disability globally. Annually,15 million stroke victims, 5 million fatalities, and an additional 5 million irreversibly damaged individuals occur globally according to the World Health Organization (WHO) [1]. A stroke has an impact on the sufferer as well as their family, job, and social surroundings. Furthermore, it can affect anyone of any age, gender, or physical condition, despite what the general public believes[2]. A stroke is an abrupt and serious neurological issue that impacts the blood vessels in the brain. It happens when there is a disturbance in the blood flow to a specific area of the brain, leading to a lack of oxygen for the brain's cells. Hemorrhagic stroke and ischemic stroke are the two forms of stroke. The former can cause mild to severe permanent damage, while the latter is rare and results from a blood artery burst that can cause brain hemorrhage. The most common type of stroke is ischemic, occurring when an artery narrows or becomes blocked, halting the flow of blood to a specific area of the brain. [3]. Particularly machine learning (ML) and artificial intelligence (AI) are quickly becoming crucial for the early diagnosis of a variety of illnesses, including diabetes.(as a classification task [4] or regression task for continuous glucose prediction [5]), hypertension [6], cholesterol [7], COVID-19 [8], sleep disorders, hepatitis C [9], CKD [10], etc. During this investigation, we will pay special attention to the stroke. Many research papers have employed machine learning algorithms to study this specific medical condition.

The techniques for developing effective binary classification models to predict the frequency of strokes are presented in this research. The synthetic minority over-sampling (SMOTE) technique was used. [11] to assess the impacts of class balance. The source of the dataset was Kaggle. Initially, the imbalanced dataset was used to construct and assess models. The effect of class balancing on model performance was subsequently demonstrated by reevaluating these models on a balanced dataset generated using SMOTE. Despite extensive research on stroke prediction, there is still a gap in analyzing model performance before and after class balancing procedures. In order to address this, this study uses SMOTE to create binary classification models on both balanced and imbalanced stroke datasets. We highlight the significance of resolving class imbalance for successful stroke prediction models by

analyzing these models in both circumstances and offering insights into how well class balancing improves model performance. We conducted an assessment of Naive Bayes, K-NN, Support Vector Machine, Random Forests, Decision Trees, and Logistic Regression. In our evaluation of both imbalanced and balanced datasets, we illustrated the efficacy of the models by achieving high AUC, precision, recall, F-measure, and accuracy. The comparative analysis emphasizes the importance of class balance for creating effective stroke prediction models.

II. LITRATURE REVIEW

Recent studies have focused on applying machine learning approaches to predict stroke risk, as part of the research community's efforts to build tools and tactics to detect and anticipate serious illnesses. Numerous research have shown how well machine learning models can predict the occurrence of strokes. For example, [12] highlighted machine learning (ML) 's crucial role in determining stroke risk variables and predicting stroke occurrences. By analyzing patient data and predicting stroke using a range of machine learning algorithms, they demonstrated the value of these models in healthcare settings.

The class imbalance in datasets, where non-stroke cases greatly outnumber stroke cases, is one of the main obstacles to stroke prediction. This disparity may cause models to become skewed in favor of the majority class, which would impair the model's predictive power for stroke incidence. Several class balancing methods have been suggested to address this issue, with the Synthetic Minority Over-sampling Technique (SMOTE) being one of them.

[13] introduced SMOTE, a technique that generates synthetic samples for the minority class by creating interpolations based on existing examples of the minority class. This method has been widely used in medical research to enhance machine learning models' performance on unbalanced datasets.

[7] investigated the use of SMOTE in COVID-19 and cholesterol prediction models. According to their findings, SMOTE successfully reduced the issue of class imbalance, improving performance indicators like accuracy, precision, recall, and F-measure.

Several machine learning models have been assessed for their efficacy in the context of stroke prediction. The techniques commonly used are naive bayes, k-nearest neighbors (KNN), support vector machines (SVM), decision trees, random forests, and logistic regression. The advantages and disadvantages of each model vary depending on the dataset and its intended use.

For example, [9] demonstrated the interpretability and simplicity of logistic regression in contrast to the complexity of decision trees by using both logistic regression and decision trees to predict hepatitis C. In a similar vein, Wang et al. (2021) showed the great accuracy and resilience of ensemble approaches like random forests by using random forests and SVMs for the prediction of chronic kidney disease.

Despite advancements, challenges remain in optimizing model performance and ensuring robustness in diverse clinical settings. Addressing class imbalance in medical datasets, which impacts predictive accuracy, is crucial. Techniques like SMOTE are used to mitigate this, but their effectiveness varies across different contexts and datasets. A key research gap is the detailed comparison of model performance before and after applying SMOTE for stroke prediction. Closing this gap is essential to validate advanced class balancing methods.

III. METHODOLOGY

A. Dataset Description

Our study was developed upon a Kaggle dataset (“Stroke Prediction Dataset,” n.d.). There were 5110 participants in total, and the attributes are all stated in Table 1 (One attribute for the target class and ten attributes as input to ML models):

TABLE I. DATASET ATTRIBUTES

Attribute name	Attribute Description
id	A distinct number assigned to every patient
gender	"Male", "Female" or "Other"
age	Patients' ages (1–82)
hypertension	0-no hypertension, 1-hypertension
heart_disease	0-no heart diseases, 1-heart disease
ever_married	"No" or "Yes"
work_type	"Children","Govt_job","Never_worked", "Private","Self-employed"
residence_type	"Rural" or "Urban"
avg_glucose_level	Blood glucose level on average
bmi(Kg/m2)	Body mass index
smoking_status	"Formerly smoked", "never smoked", "smokes", "Unknown"
stroke	1-Stroke, 0- no stroke

(*Source: Author's compilation)*

There are far more non-stroke cases in the original dataset than stroke patients, resulting in an imbalance. It's given in Figure 1.

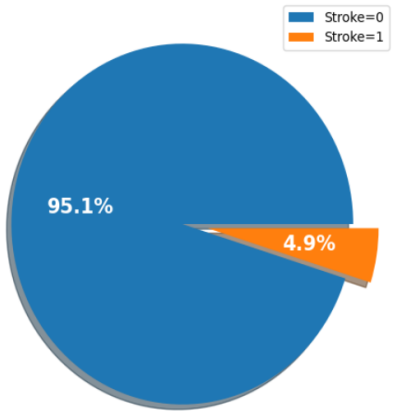


Fig 1-Distribution of Stroke Before apply SMOTE

(Source: Author's compilation)

B. Data Preprocessing

There are 5,110 items in the collection. To preserve data integrity and avoid analytical distortion, careful data preprocessing was essential, including filling in missing values and eliminating outliers. These procedures guaranteed a dependable and clean dataset for additional research and model training. We employed the Synthetic Minority Over-sampling Technique (SMOTE) to tackle the imbalance in the class, generating synthetic samples for the minority class consisting of stroke sufferers. By balancing the dataset, improves the models' capacity to learn and provide precise stroke predictions. Figure 2 - A displays a balanced distribution of the target variable, stroke, after applying SMOTE. This illustrates how SMOTE can effectively solve class imbalance by guaranteeing that machine learning models can predict stroke occurrences properly and impartially toward the majority class. A balanced dataset's age distribution is depicted in Figure 2 - B, with the majority of stroke victims falling into the 70–79 age range, then the 50–59 age group. This emphasizes the necessity of providing older persons with specialized healthcare resources and preventative measures. The gender distribution in the balanced dataset is displayed in Figure 2 - C. Males are more likely than females to experience strokes, with most ladies never experiencing one. This distribution draws attention to gender disparities in the incidence of stroke and implies that these variations may need to be taken into account in targeted interventions.

Figures 2 - D and 3 - A below show heart disease and hypertension prevalence in stroke participants. We note that a significant proportion of stroke survivors in both figures do not have a diagnosis of heart disease or hypertension.As seen in Figure 3 - B, the majority of stroke patients are overweight. This graphic highlights the significance of tracking and controlling weight as a preventive strategy against stroke by suggesting a possible link between increased body weight and the risk of stroke.Figure 3 - C shows that most stroke victims work in private employment. This shows that those in the private sector may have a greater incidence of stroke,

potentially as a result of lifestyle choices connected to job and stress. This realization highlights the necessity of workplace prevention and focused health interventions.

To sum up, the visual aids provide a comprehensive view of stroke distribution based on various characteristics. Key findings include a higher incidence of stroke among older age groups, particularly those aged 70-79, and a noticeable gender disparity, with more women affected than men. Additionally, individuals with private employment and those who are overweight show higher stroke risk. These insights underscore the need for targeted preventive measures and health interventions for at-risk populations, crucial for effective stroke prevention and management strategies.

C. Machine Learning Models

The classifiers utilized in this portion include Naive Bayes, Logistic Regression, Decision Trees, Random Forests, SVM, and KNN for the purpose of predicting strokes.

In our system, the first model that is used is logistic regression (LR) [14]. Logistic regression, initially created for binary tasks, is a statistical technique used for data classification and has since been expanded to handle multi-class issues. The outcome from the model is a binary variable denoted as p=P(Y=1), representing the probability of an instance being in the "Stroke" class. Conversely, 1−p=P(Y=0) represents the likelihood of an instance belonging to the "Non-Stroke" class.The following is the expression for the relationship between the log-odds and the model parameters β_i:

$$log\left(\frac{p}{1-p}\right) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_n X_n$$

Next we used Decision Trees, which logically and clearly classify data by splitting it into branches based on feature values. Then we employed an ensemble of decision trees called a random forest classifier after the decision tree. Random forests lower the chance of overfitting and increase accuracy by averaging the predictions of several trees. Then we used naive Bayes classifier assumes predictor independence based on Bayes' theorem. Despite its simplicity, it processes huge datasets efficiently and produces accurate results for classification tasks. When classifying a new subject i with characteristics vector f_i the maximum value of $P(c|\ f_{i1},\ldots,\ f_{in})$ corresponds to class c. The conditional probability is defined in the following way:

$$P(c|\ fi1,\ldots,fin)\ = \frac{P(c|\ fi1,\ldots,fin)\ Pc()}{P(\ fi1,\ldots,fin)}$$

Next, the K-nearest neighbors (KNN) classifier uses distance-based methods, such as Manhattan or Euclidean distance, to determine how similar or different two examples are from one another. For a new sample, the KNN finds the nearest K vectors and classifies the sample based on the majority class of its neighbors, aiding in stroke prediction [15]. A support vector machine (SVM) model was also used, which finds the optimal hyperplane with the biggest margin for splitting the data points into different classes. SVMs are flexible with various kernel functions and efficient in high-dimensional spaces.

D. Evaluation

Several performance measures were noted during the ML models under consideration's evaluation procedure. We shall take into account the most widely utilized in the pertinent literature in the current analysis [16].
Recall is the percentage of stroke survivors who, in relation to all positive participants, were correctly classified as positive. Sometimes it's referred to as true positive rate or sensitivity. Precision and recall are better suitable for detecting a model's flaws when dealing with unbalanced data. The accuracy shows how many actual stroke sufferers fit into this group. The recall rate is the percentage of stroke patients who can be anticipated with accuracy. The F1-Score, which offers a summary of a model's prediction ability, is the harmonic mean of precision and recall.

$$\text{Recall} = \frac{\text{TP}}{\text{TP + FN}'} \qquad \text{Precision} = \frac{\text{TP}}{\text{TP + FP}}$$

$$\text{F1-Score} = 2.\frac{\text{Precision . Recall}}{\text{Precision + Recall}'} \qquad \text{Accuracy} = \frac{\text{TN + TP}}{\text{TN+TP+FN+FP}}$$

Please bear in mind that FP represents false positive, FN represents false negative, TP represents true positive, and TN represents true negative.

IV. RESULTS AND DISCUSSION

In this segment, we compare the results of six different machine learning models—Naive Bayes, Logistic Regression, Random Forests, Decision Trees, SVM, and KNN—both prior to and after implementing SMOTE for predicting stroke risk. The

performance of each model was assessed using accuracy, precision, recall, and F1-score to obtain a complete overview, accounting for class imbalance in the analysis. Table 2 summarizes the performance measures on the unbalanced dataset, while Table 3 presents the metrics after balancing the dataset with SMOTE.

TABLE II. MODEL PERFORMANCE BEFORE SOMTE

ML Model	Precision	Recall	F1-Score	Accuracy
LR	0.92	0.96	0.94	0.96
DT	0.93	0.93	0.93	0.93
RF	0.92	0.96	0.94	0.96
NB	0.94	0.89	0.91	0.89
K-NN	0.96	0.96	0.94	0.96
SVM	0.92	0.96	0.94	0.96

TABLE III. MODEL PERFORMANCE AFTER SMOTE

ML Model	Precision	Recall	F1-Score	Accuracy
LR	0.94	0.74	0.82	0.74
DT	0.93	0.88	0.91	0.88
RF	0.93	0.92	0.92	0.92
NB	0.95	0.91	0.73	0.61
K-NN	0.93	0.8	0.85	0.8
SVM	0.96	0.61	0.72	0.61

(*Source: Author's compilation)*

A. Model Performance Visualization

The effect of correcting class imbalance on accuracy, precision, recall, and F1-score is illustrated through visual depictions of model performance before and after SMOTE. Figure 4-A shows the accuracy of each model, highlighting performance enhancement with balanced data. Figure 4-B compares precision before and after SMOTE, indicating increased precision in stroke prediction. Figure 4-C illustrates recall metrics, showing improved ability to correctly identify stroke patients after addressing class imbalance. Figure 4-D presents the F1-score, combining recall and precision, offering an overview of overall model performance improvement with SMOTE. Models with the highest accuracy before SMOTE are Random Forest, KNN, SVM, and Logistic Regression, each achieving about 96%, while Naive Bayes has the lowest at 89%. After SMOTE, Random Forest retains the highest accuracy at roughly 92%, while Naive Bayes and SVM drop to around 61%. These visualizations and insights highlight the impact of correcting class imbalance on model performance, showing that balancing the dataset with SMOTE enhances performance to varying degrees, with some models benefiting more significantly than others. This demonstrates the importance of addressing class imbalance in improving predictive accuracy and reliability.

V. CONCLUSION

In this work, we examined how class balance affects the predictive power of many machine learning models in stroke incidence using data from Kaggle. The study consisted of two phases: modeling the original imbalanced dataset and modeling a balanced dataset created using the Synthetic Minority Over-sampling Technique (SMOTE). Initial evaluations on the imbalanced dataset showed high accuracy (around 96%) for models like Logistic Regression, Random Forest, KNN, and SVM, while Naive Bayes lagged at 89%. The problem of class imbalance has been highlighted by this difference, where the overwhelming presence of the majority class can give the impression of models being highly accurate. The models' performance metrics saw a substantial alteration following the use of SMOTE. The Random Forest model demonstrated the highest accuracy at 92%, indicating its robustness. Conversely, the accuracy of models like SVM and Naive Bayes dropped to around 61%, showing their sensitivity to changes in data distribution. The dataset balancing generally results in more realistic performance evaluations, as indicated by the comparison of precision, recall, F1-score, and accuracy before and after SMOTE. This is crucial for medical diagnostics,

where false negatives can have serious consequences, such as failing to predict a stroke. The analysis reveals that no single model excels across all metrics. The choice of a model should take into account the particular needs and preferences of the application. For example, following the deployment of SMOTE, Random Forest continued to perform well across a number of metrics. Future research should explore hybrid approaches that combine several models and include more feature engineering and selection strategies to improve predictive performance. Furthermore, putting these models to the test in actual clinical situations would shed light on how reliable and useful they are. To sum up, resolving class imbalance through methods such as SMOTE is essential to building strong and effective machine learning models for stroke prediction. This research underscores the importance of this step in ensuring that models provide accurate and balanced predictions, contributing to early and accurate stroke detection and timely intervention.

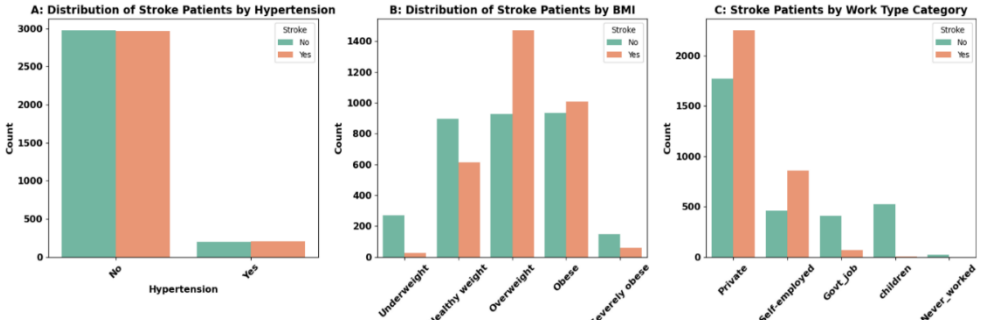


Fig 3 -Hypertension, BMI, Work Type distributions *(Source: Author's compilation)*

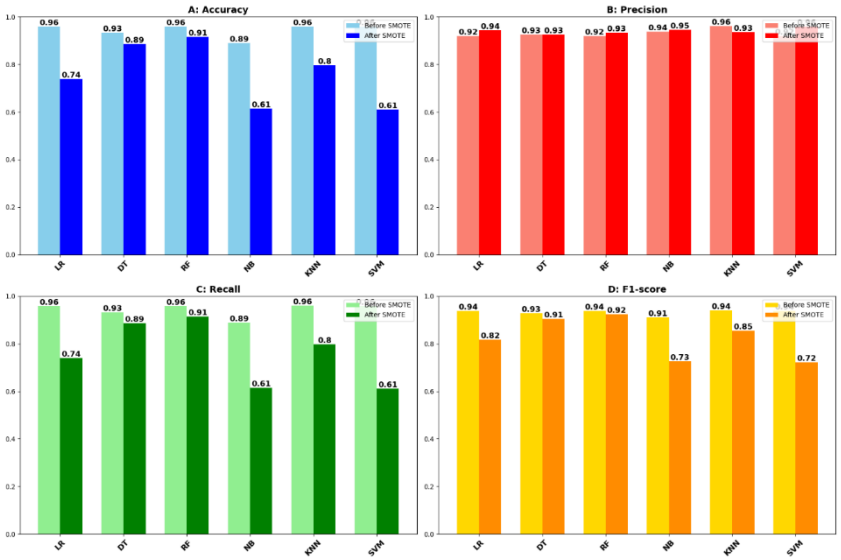


Fig 4-Accuracy, Precision, Recall, F1-Score comparison (*Source: Author's compilation*

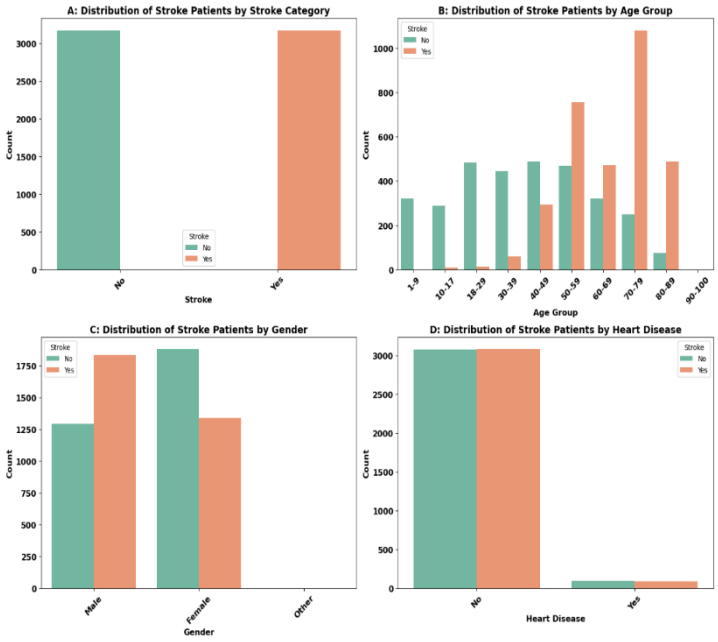


Fig 2 - Balanced dataset, Age, Gender, Heart distributions (Source: Author's compilation)

REFERENCES

[1] "WHO EMRO | Stroke, Cerebrovascular accident | Health topics," World Health Organization - Regional Office for the Eastern Mediterranean. Accessed: Jun. 02, 2024. [Online]. Available: http://www.emro.who.int/health-topics/stroke-cerebrovascular-accident/index.html

[2] T. Elloker and A. J. Rhoda, "The relationship between social support and participation in stroke: A systematic review," *Afr J Disabil*, vol. 7, p. 357, 2018, doi: 10.4102/ajod.v7i0.357.

[3] A. Bustamante *et al.*, "Blood Biomarkers to Differentiate Ischemic and Hemorrhagic Strokes," *Neurology*, vol. 96, no. 15, pp. e1928–e1939, Apr. 2021, doi: 10.1212/WNL.0000000000011742.

[4] N. Fazakis, O. Kocsis, E. Dritsas, S. Alexiou, N. Fakotakis, and K. Moustakas, "Machine Learning Tools for Long-Term Type 2 Diabetes Risk Prediction," *IEEE Access*, vol. 9, pp. 103737–103757, 2021, doi: 10.1109/ACCESS.2021.3098691.

[5] S. Alexiou, E. Dritsas, O. Kocsis, K. Moustakas, and N. Fakotakis, *An approach for Personalized Continuous Glucose Prediction with Regression Trees*. 2021, p. 6. doi: 10.1109/SEEDA-CECNSM53056.2021.9566278.

[6] E. Dritsas, N. Fazakis, O. Kocsis, N. Fakotakis, and K. Moustakas, "Long-Term Hypertension Risk Prediction with ML Techniques in ELSA Database," 2021, pp. 113–120. doi: 10.1007/978-3-030-92121-7_9.

[7] N. Fazakis, E. Dritsas, O. Kocsis, N. Fakotakis, and K. Moustakas, *Long-term Cholesterol Risk Prediction using Machine Learning Techniques in ELSA Database*. 2021, p. 450. doi: 10.5220/0010727200003063.

[8] A. S. Kwekha-Rashid, H. N. Abduljabbar, and B. Alhayani, "Coronavirus disease (COVID-19) cases analysis using machine-learning applications," *Appl Nanosci*, vol. 13, no. 3, pp. 2013–2025, Mar. 2023, doi: 10.1007/s13204-021-01868-7.

[9] M. A. Konerman *et al.*, "Machine learning models to predict disease progression among veterans with hepatitis C virus," *PLOS ONE*, vol. 14, no. 1, p. e0208141, Jan. 2019, doi: 10.1371/journal.pone.0208141.

[10] W. Wang, G. Chakraborty, and B. Chakraborty, "Predicting the Risk of Chronic Kidney Disease (CKD) Using Machine Learning Algorithm," *Applied Sciences*, vol. 11, no. 1, Art. no. 1, Jan. 2021, doi: 10.3390/app11010202.

[11] S. Maldonado, J. López, and C. Vairetti, "An alternative SMOTE oversampling strategy for high-dimensional datasets," *Applied Soft Computing*, vol. 76, pp. 380–389, Mar. 2019, doi: 10.1016/j.asoc.2018.12.024.

[12] E. M. Alanazi, A. Abdou, and J. Luo, "Predicting Risk of Stroke From Lab Tests Using Machine Learning Algorithms: Development and Evaluation of Prediction Models," *JMIR Form Res*, vol. 5, no. 12, p. e23440, Dec. 2021, doi: 10.2196/23440.

[13] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, "SMOTE: Synthetic Minority Over-sampling Technique," *Journal of Artificial Intelligence Research*, vol. 16, pp. 321–357, Jun. 2002, doi: 10.1613/jair.953.

[14] S. Nusinovici *et al.*, "Logistic regression was as good as machine learning for predicting major chronic diseases," *Journal of Clinical Epidemiology*, vol. 122, pp. 56–69, Jun. 2020, doi: 10.1016/j.jclinepi.2020.03.002.

[15] P. Cunningham and S. J. Delany, "k-Nearest Neighbour Classifiers - A Tutorial," *ACM Comput. Surv.*, vol. 54, no. 6, pp. 1–25, Jul. 2022, doi: 10.1145/3459665.

[16] H. M and S. M.N, "A Review on Evaluation Metrics for Data Classification Evaluations," *IJDKP*, vol. 5, no. 2, pp. 01–11, Mar. 2015, doi: 10.5121/ijdkp.2015.5201.

Enhancing Personalized Tourism Recommendations in Sri Lanka: Integrating Sentiment Analysis and Semantic Clustering for Tailored Travel Experience

N.G.D. Nethmini
Department of Computer and Data Science
National School of Business Management
Homagama, Sri Lanka
ngdnethmini@students.nsbm.ac.lk

S.A.D.H.M. Samarathunga
Department of Computer and Data Science
National School of Business Management
Homagama, Sri Lanka
dhmsamarathunga@students.nsbm.ac.lk

A.S.A. Gunathilaka
Department of Computer and Data Science
National School of Business Management
Homagama, Sri Lanka
asagunathilaka@students.nsbm.ac.lk

Ms. Lakni Peiris
Department of Software Engineering and Information Systems
National School of Business Management
Homagama, Sri Lanka
lakni.p@nsbm.ac.lk

***Abstract*— One of the major challenges faced by Sri Lanka's tourism industry is to provide personalized travel experiences for global visitors. Sentiment analysis and semantic clustering-based novel systems have been introduced in this research paper to overcome these gaps. This system analyzes extensive amounts of tourists' reviews that give important insights into cognitive behaviors as well as emotional tones, making it possible to provide appropriate personalized context-sensitive travel advice. Advanced natural language processing techniques involving TF-IDF Vectorizer and K-means clustering have been used to organize reviews into meaningful clusters. This reveals hidden tourist spots that would otherwise go unnoticed. Further, the recommendation system has employed Bayesian classification, big data analytics, and machine learning to improve accuracy and scalability. The aim of developing this system is to address the differences between tourists' preferences and available places in Sri Lanka. Promoting such lesser-explored destinations and offering tailored experiences improves international travelers' overall satisfaction. Such a development can be a turning point for the Sri Lankan tourism industry since it advances its sustainability through growth, attracting new client bases.**

Keywords—Semantic analysis, semantic clustering, personalized recommendation system, tailored travel experience, natural language preprocessing

I. INTRODUCTION

Tourism has been involved dynamically in economic, cultural, and social sectors by thriving countries worldwide. Engaging in an enormous number of activities, leisure, and travel related to health, business, etc., people pay attention to travel to foreign countries. Among other foreign countries, Sri Lanka represents a considerable number of tourist attractions as a result of highlighting statistics from the UNWTO, with 1.49 million international tourist arrivals in 2023, with 3829 establishments available for tourists for accommodation facilities. The country is a tropical island in the Indian Ocean located between the sea routes of Europe and the Far East [3]. The geographical location implied the colonization by three Western powers. Sri Lanka offers a wide range of tourist attractions for foreigners: beach destinations, cultural heritage, favorable climate, national parks, wildlife, and religious locations. Expounding the seasonality of Sri Lanka, January to March and December are peak seasons where July, August, and November cluster as mini-peak months for tourists' arrivals, where the greatest number of tourists recorded in December. The constraints and challenges that impact tourism in Sri Lanka, lack of high-quality accommodation is a concern as the SLTDA annual report indicates classified tourist hotels limited to 16 686, which is identified as a concern when compared to Singapore and Thailand [3]. The second fact is the storage of trained workers to provide standard hospitality services. Attracting quality tourists has been a challenge due to the infrastructure and support of the system. The spotlight of the challenges narrows down to the consequences of war, COVID-19, and the economic crisis in Sri Lanka.

Tourists' motivation is influenced by categories that may arise from deep-rooted and psychological needs. The mentioned categories are physical motivation to gain refreshment for body and mind, curiosity about

foreign countries to experience traditional events, food culture, history and arts, bond with new individuals to interact with international society, and the ultimate category noted as prestige and status which defines engaging in hobbies, continuation of education, following business goals [2]. The mentioned psychological motives impact the human mindset to travel to foreign countries thus, tourists concentrate on facts regarding places and their present status based on economic, environmental, social, cultural, etc. The system is dedicated to foreign tourists to explore places that match their motives and influence their mindset. The cause of the research follows the identified problem as international tourists struggle to identify hidden tourist attraction places that meet their preferences and motives due to unstructured data online.

The research problem stated that lack of attention and approach of psychological fact drive towards unable to capture the foreign tourists' cognitive behavior and mindset. Analyzing the past reviews of foreign tourists about Sri Lankan tourist attraction places utilizing natural language processing, sentimental analysis, semantic clustering, and data visualization gain valuable business insights to the target population [2],[5],[7],[16].

Existing systems lack analyzing tourists' thoughts and interpreting sentimental and semantic changes by utilizing user-generated content. Consequently, system outcomes turn out to carry contradictions between tourists' motives and pre-identified results. The proposed research initiates a novel approach to uplift the tourism industry in Sri Lanka by developing a model that influences sentimental analysis and sentimental clustering for foreign tourists. The research institute on merging notions such as natural language processing, sentimental clustering, semantic analysis, and explainable AI to generate highly personalized and context-aware recommendations for tourists.

The research paper enhances the following areas and aspects. Limited personalization in current systems points out the requirements of retrieving accurate, personalized results according to user queries. In the last two decades, the search method has been prioritized to match users' desired inputs in the system. Search queries and algorithms are generalized due to a lack of personalization. Concerning high dimensionality to boost a wide range of contents (choices) for users able to level up the user interaction and engagement through personalization [4]. Uplifting user experience and interaction influences the recommendation system.

Addressing contextual factors such as weather, time of day, and user preferences is crucial for delivering a truly personalized experience. Current systems often lack the provision of summaries of recommended travel spots with native language support, which can significantly enhance the overall user experience [5]. Moreover, cross-cultural recommendations are vital for tourists visiting Sri Lanka, and systems should be designed to accommodate these cultural differences [5]. There is a need for systems capable of handling large volumes of data while ensuring robust user privacy protections [6]. Integration with other services, such as transportation and accommodation providers, remains underexplored but has the potential to enhance the overall tourist experience significantly [7]. Evaluation metrics such as precision, recall, and F1-score are essential for assessing the performance of personalized recommendation systems, yet they are often overlooked in existing research [8]. Furthermore, emerging technologies like artificial intelligence and machine learning offer significant advancements for these systems.

II. LITERATURE REVIEW

Tourism has become vital in impacting economic, cultural, and social dimensions worldwide. Integrating personalized recommendation systems using sentimental analysis and sentimental clustering for foreign tourists enhances the experiences they receive and is an uplifting development for the tourism industry. This is especially a significant step for Sri Lanka, which is known for its diverse tourist attractions yet has minimal options of recommendations and knowledge for tourists. Through the use of such techniques, Sri Lanka can offer tourists personalized recommendations to ensure satisfaction and ease in increasing tourism within the country.

Sri Lanka offers a rich cultural heritage, beautiful landscape, diverse wildlife, and many more, such as UNESCO world heritage sites like Sigiriya and Kandy and a lot more diverse scenarios and festivals. Despite these attractions, many unique destinations may remain underexplored to many tourists who are not familiar with Sri Lanka. Implementation of personalized recommendation systems can address this by analyzing reviews to provide tailored recommendations from a wide range of tourists, promoting lesser-known areas to more tourists. By showcasing these hidden gems, Sri Lanka can enhance the satisfaction of tourists, ensure a greater distribution of tourism around the island, and boost the overall growth and sustainability of the tourism sector.

A. AI technologies in tourism

1) Sentiment analysis

This AI technique interprets the emotional tone of text from data to help understand the preferences and emotions of the data. Current tourism

recommendation systems apply sentiment analysis and use semantic clustering, which helps to organize and analyze the tourist feedback, which will help understand them better [7]. This can be used to figure out and examine the hidden emotions behind different tourist reviews and opinions to help tailor recommendations to be more accurate and up to expectations.

2) Bayesian classification

A statistical method based on Bayes theorem happens to be an AI technique that significantly influences many applications and systems in tourism researchers. This classification technique is most used in predictive analysis to make predictions based on probabilities using past data. Bayesian classifiers have been used to forecast tourist arrival timelines, categorize their preferences, and segment and cluster markets based on some attributes [3].
This classification technique can be used to update the prediction system when new data are created, making it possible to be updated and adaptable to any change in the trend of the tourism industry [1]. Initially, they were used to model and predict tourism behavior under factors like economic condition changes or travel restrictions. But now they can assist in making policies and even the business sector to help make prior marketing strategies and better decisions to better cater to tourists [1].

3) Big data analytics

Big data analytics play a vital role in collecting and analyzing vast amounts of data. By processing data such as reviews from other tourists, the tourism stakeholder will gain valuable insights into the behaviors, mindsets, and preferences of tourists. According to the study by [17], big data is essential as it enables the extraction of actionable insights from large datasets, which can be used to influence and leveraged into creating more relevant recommendations. In Sri Lanka, too, this concept can be used to uncover these precious insights.

4) Machine learning

Machine learning algorithms and techniques are also vital for developing these sophisticated recommendation systems. This allows us to analyze historical data and predict future behaviors and preferences for customized suggestions. Clustering algorithms group tourists with similar interests and recommend destinations and activities to align with user preferences [8]. Similarly, in Sri Lanka, machine learning can further help identify and promote lesser-known attractions to new tourists.

5) NLP (natural language processing)

NLP is further an essential concept for sentiment analysis as it involves inspecting textual reviews from tourists to assess and evaluate their sentiments and opinions on different destinations and activities. According to [18], sentiment analysis is vital in examining tourist experiences to refine recommendations.

6) Applications in tourism

a) Global

Globally, many studies and research show the highly effective aspect of personalized recommendation systems in tourism. [19] It is also seen that personalized recommendations increase tourists' satisfaction. Similarly, the recent advancements in sentiment analysis and semantic clustering have been showcased in global case studies, showing how effective tailoring tourism experiences can be and also understanding and addressing the diverse needs of tourists [15].

b) Locally (Srilanka)

In the context of Sri Lanka, there is an increasing interest in the idea of personalized recommendation systems to enhance tourism. Certain local studies have explored the use of sentiment analysis to get tourist feedback for better services and recommendations. As referred to by [16], it is discussed the potential of data-driven approaches in personalizing tourist experiences. This study highlights how data from various sources like social media, travel blogs, and others can be analyzed to extract insights about tourist preferences.

7) Technological integration

Integrating advanced technologies like big data analytics, machine learning, and natural language preprocessing (NLP) into the sector of tourism further enhances the advantages present for tourists.

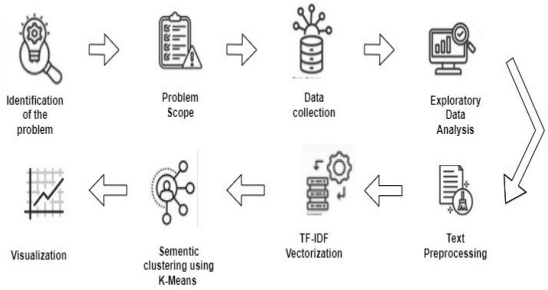


Fig 1. Methodology

III. METHODOLOGY

A. Data collection

The research problem addressed identifying psychological behavior patterns by analyzing past tourists' reviews to utilize natural language processing, semantic clustering, sentimental analysis, and data visualization concepts. At the initial stage, known as data collection, the secondary method was selected to capture a large volume of data; therefore, an appropriate dataset was selected for carrying out the purpose of a quantitative study. From the data preprocessing step, a randomly selected sample from the dataset was derived to proceed with the analysis and clustering algorithms.

The dataset used for the research consists of a compilation of tourist reviews obtained from Mendeley Data [20]. It consists of 16,000 reviews from foreign tourists who have visited beaches, wildlife parks, and cultural and religious sites in Sri Lanka. The dataset was gathered between 2010 and 2023. The dataset record features a review ID, a section with unidentifiable characteristics, age of the person, nationality, and purpose of travel, the review, and a single score from '1' to '5', the rating being how satisfied the tourist was with the visit. Using this dataset, the study intends to obtain scattered patterns and insights focused on tourist behavior and preferences and enhance the recommendation system.

B. Data preprocessing

Consequently, raw data should be cleansed in a proper data preprocessing procedure. Obtaining the overall understanding of variables, types, and statistical interpretations following text preprocessing steps executed aligning the natural language processing concepts. The lowercasing step was utilized to emphasize the uniformity in the text to normalize the serve in between capitalizing vs simple words. Pandas and NumPy libraries are applied to access flexible data structures and vectorized purposes. The next step was removing the punctuation to assist the process of tokenization. Spelling correction is concerned with deriving meaningful notions utilizing the Text Blob library. Removing the stop words step used carrying the affection of a small scale of semantic weight and clutter analysis. Examples of articles, conjunctions, and prepositions are included. Nltk. Corpus, known as the natural language toolkit library, is approached to obtain a considerable range of text corpora and lexical resources. Moving towards a salient step in NLP tokenization refers to splitting text into smaller units named tokens, accomplishing the task of transforming data into a structured format for analysis purposes through NLTK. The stemming step included reducing words to their base or root form. Pointing an example as "runner," "running," and "ran" are all reduced to "run" through NLTK. With the previous step, lemmatization is utilized to reduce words to their base or root form by considering the context and morphological analysis of words. Engaging in natural language processing, cleaning, and preparing raw reviews executed for future semantic analysis purposes.

C. Feature extraction

TF-IDF Vectorizer was employed to convert the preprocessed data into a numerical representation, capturing the importance of each feature in the dataset. This technique assigns weights to words based on their frequency and rarity across the entire corpus, enabling the identification of distinctive features that characterize user preferences.

The formula for TF-IDF is:

$$TF\text{-}IDF(t, d) = TF(t, d) \times IDF(t) \quad [1]$$

Where*:*

• TF (t, d) is the term frequency of term t in document d.

• IDF(t) is the inverse document frequency of term t, calculated as log (N / | {d ∈ D: t ∈ d}|) where N is the total number of documents and |{d ∈ D : t ∈ d}| is the number of documents containing term t.

The Bag of Words model was used to represent each user and item (attraction, hotel, or restaurant) as a bag, or a set, of its constituent words. This approach enables the calculation of similarities between users and items based on their word frequencies, facilitating personalized recommendations. In this model, each text is represented as a vector of word counts without considering word order or grammar.

The integration of TF-IDF Vectorizer and Bag of Words in the recommendation system enables the identification of subtle patterns and relationships in user behavior, leading to more accurate and personalized recommendations.

D. Semantic Clustering

Semantic clustering refers to a procedure of grouping clusters according to the meaning of the texts. Semantic clustering merges with the concept of natural language processing and data analysis. Under the research, semantic clustering was used to analyze the tourists' reviews of the tourists' places in Sri Lanka by grouping similar texts that contain similar meanings. In the clustering procedure, a subset of attributes is introduced under the target object. Identifying a specific name for the root cluster, the mentioned attribute subsets will be called after the cluster instance [9]. For the research, in semantic analysis, k means clustering has been approached under an unsupervised learning method. To initial the k mean clustering, the number of clusters which refer as the k value. To identify the value, the entropy concept is feasible to utilize as follows through the below equations [10].

$$JKM1 = \Sigma\, i=1\ n\ [x]k = lcxik \,||\, xi - ak \,||\, 2 + [x]k = lclk \quad (1)$$

[2]

$$JK\text{-}means = \Sigma\, i=1\ n\ [x]k = lcxik \,||\, xi - ak \,||\, 2 + [x]k = lclk \quad \Sigma\ i=1\ n\ [x]k = lclk \quad (2)$$

[3]

The K-means clustering for the semantic analysis was conducted using the sklearn library by importing K-means and silhouette scores. Furthermore, matplotlib utilized cluster boundaries for the purpose as to obtain an understanding of the performance of the K-means algorithm. From the analysis perspective, values of K are derived using the number of classes. A data clustering algorithm can be used as a classifier by approaching omitted class attributes to assess the cluster results [11]. Therefore, the semantic clustering method narrows down to supervised learning. Applying both methods efficiently, for the research, the K value was determined for 5. Subsequently, cluster labels will be defined for each review in the dataset. Identified cluster labels will be attached to the data frame. Once the algorithms are executed, the silhouette score will be checked to identify the quality of the cluster. The below performance factors were approached to evaluate the quality of the output. The initial stage reaches with accuracy to measure the level of predicting clusters to actual clusters. The accuracy rate represents how clusters have predicted accurately to match real semantic categories of the reviews. Confusion matrix impacts to evaluation cluster performance by analyzing the volume of reviews from each true category allocated for predicted clusters. The ROC curve plots the true positive rate against the false positive rate for each class, and the Area Under Curve figures out the model's ability to distinguish between classes. The below formulas examine how cluster performance evaluates utilizing mentioned factors.

IV. RESULTS AND DISCUSSION

Pursuing the research objective of identifying the tourist's destinations, ideas, and activities through analyzing foreigners 'cognitive behavior, thoughts & ideas in a verbal manner by utilizing natural language processing, semantic clustering, and analysis, the following observations were derived. According to the methodology process, the steps were conducted for the random sample and the whole dataset. By deriving the k value through the elbow method, for the sample data, the k value was 2, and for the whole dataset, 5. After deriving the k value, the silhouette score was calculated for the sample as well as for the entire dataset. Utilizing data visualization, the outcomes of the semantic clustering plotted in the scatter plot represented 5 clusters for the whole dataset in Fig 2 (clusters: beaches -0, agricultural and village adventures- 1, nature and wildlife-2, historic - 3, religious – 4) Carrying out the purpose of evaluating the clusters, ROC curve, AUC, and confusion matrix performance factors were executed (Fig 3). Based on the results of the ROC curve, classes 0,1, and 4 consist of 0.72 AUC value, which indicates a considerably good level of performance, while classes 2 and 3 performed weakly. ROC curve interpretation narrows down to complex situations as it's unable to define the "positive" and "negative" hence, AUC values interpret the measures for the overall performance of the model. The confusion matrix was able to derive the values of the number of correct assignments to the clusters as 1690, 1595, 1197,1529,1358. The higher number of incorrectly predicted classes is under cluster 02 as it defines nature as the reason in most of the reviews, the content, ideas, and thoughts related to nature have given a major input. Therefore, the incorrectly predicted results have occurred due to the meaning of the content as possibilities of containing similar characteristics, which caused the clustering model to group them incorrectly. To minimize the errors and maximize the accuracy, methods such as tuning the clusters and comparing the other clustering techniques. Moving toward future considerations, developing a recommendation system for tourists

based on semantic clustering will help enhance the user experience by personalized and tailored travel suggestions, suggesting different destinations and available activities for tourists.

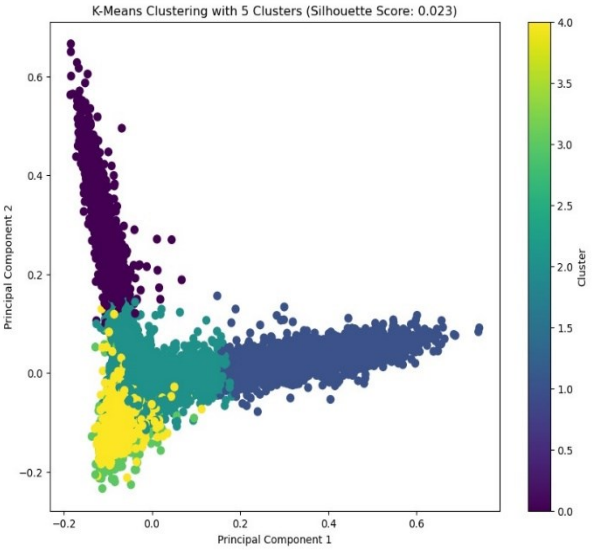


Fig 2.K-means clusters

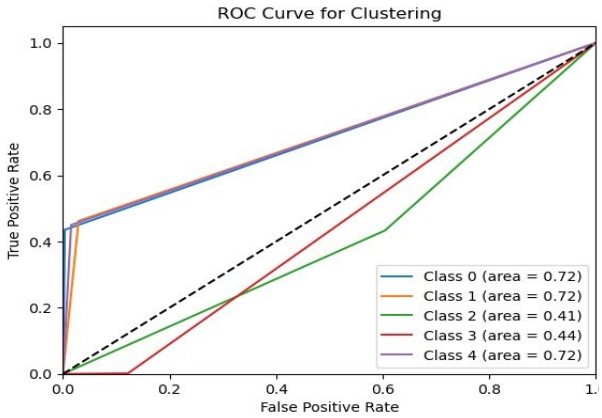


Fig 3.ROC curve for clustering

CONCLUSION

In conclusion, using sentiment analysis and semantic clustering in the sector of tourism greatly helps to focus on areas that lack attention in existing systems and related applications. Carrying out the primary objective of addressing psychological facts, the mindset of foreigners regarding tourism in Sri Lanka was essentially considered when modeling the semantic clusters through utilizing natural language processing and semantic clustering. This greatly enhances the travel experience in Sri Lanka. By using advanced AI techniques such as natural language preprocessing, sentiment analysis, semantic clustering, and data visualization, it is possible to address the gap by tailoring suggestions and the challenges of unstructured data. In this research, the aim was to help uncover hidden tourist destinations and improve the suggestions available for tourism. The research's focus is on highlighting the importance of continuously adapting and improving the recommendations that align with tourists' diverse preferences and nature. This will enhance tourism and contribute to the country as a whole.

REFERENCES

[1] Perera, K.K., 2017. An analysis of recent trends in tourist arrivals in Sri Lanka. *Athens Journal of Tourism*, *4*(1), pp.51-80.

[2] Camilleri, M.A. and Camilleri, M.A., 2018. *The tourism industry: An overview* (pp. 3-27). Springer International Publishing.

[3] Fernando, S., Bandara, J.S. and Smith, C., 2016. Tourism in Sri Lanka. In *The Routledge Handbook of Tourism in Asia* (pp. 271-284). Routledge.

[4] Modarresi, K., 2016. Recommendation system based on complete personalization. *Procedia Computer Science*, *80*, pp.2190-2204.

[5] Shiranthika, C., Premakumara, N., Weerawarnakula, J.P., Lakma, H., Fernando, S. and Sumathipala, S., 2018. Personalized Travel Spot Recommendation and Guidance System for Sri Lankan Tourists.

[6] Boralessa, G.W.Y.S., Ragel, R.G., Shafana, M.S. and MJ, A.S., 2022. Navigating the Landscape of Recommender Systems: An Extensive Literature Review. (Boralessa, 2022)

[7] Abbasi-Moud, Z., Vahdat-Nejad, H. and Sadri, J., 2021. Tourism recommendation system based on semantic clustering and sentiment analysis. *Expert Systems with Applications*, *167*, p.114324.

[8] Núñez, J.C.S., Gómez-Pulido, J.A. and Ramírez, R.R., 2024. Machine learning applied to tourism: A systematic review. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, p.e1549.

[9] Shannon, K. and Snodgrass, R.T., 1990, September. Semantic Clustering. In *POS* (pp. 389-402).

[10] Sinaga, K.P. and Yang, M.S., 2020. Unsupervised K-means clustering algorithm. *IEEE access*, *8*, pp.80716-80727.

[11] Pham, D.T., Dimov, S.S. and Nguyen, C.D., 2005. Selection of K in K-means clustering. *Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science*, *219*(1), pp.103-119.

[12] Li, J., Xu, L., Tang, L., Wang, S. and Li, L., 2018. Big data in tourism research: A literature review. *Tourism management*, *68*, pp.301-323.

[13] Ricci, F., 2010. Mobile recommender systems. *Information Technology & Tourism*, *12*(3), pp.205-231.

[14] Rehman Khan, H.U., Lim, C.K., Ahmed, M.F., Tan, K.L. and Bin Mokhtar, M., 2021. Systematic review of contextual suggestion and recommendation systems for sustainable e-tourism. *Sustainability*, *13*(15), p.8141.

[15] Shafqat, W. and Byun, Y.C., 2019. A recommendation mechanism for under-emphasized tourist spots using topic modeling and sentiment analysis. *Sustainability*, *12*(1), p.320.

[16] Nethmin, N.A.L., Saranga, L.G.S., Dilshan, A.D.T., Wijewickrama, H.E., Swarnakantha, N.R.S. and Rajapaksha, U.S., 2023, December. Enhancing Tourism Experience in Sri Lanka through Data Driven Recommendations. In *2023 International Conference on Innovative*

Computing, Intelligent Communication and Smart Electrical Systems (ICSES) (pp. 1-7). IEEE.

[17] Gretzel, U., Sigala, M., Xiang, Z. and Koo, C., 2015. Smart tourism: foundations and developments. *Electronic markets*, *25*, pp.179-188.

[18] Álvarez-Carmona, M.Á., Aranda, R., Rodríguez-Gonzalez, A.Y., Fajardo-Delgado, D., Sánchez, M.G., Pérez-Espinosa, H., Martínez-Miranda, J., Guerrero-Rodríguez, R., Bustio-Martínez, L. and Díaz-Pacheco, Á., 2022. Natural language processing applied to tourism research: A systematic review and future research directions. *Journal of king Saud university-computer and information sciences*, *34*(10), pp.10125-10144.

[19] Gavalas, D., Konstantopoulos, C., Mastakas, K. and Pantziou, G., 2014. Mobile recommender systems in tourism. *Journal of Network and computer applications*, *39*, pp.319-333.

[20] Sewwandi, Taniya (2023), "Tourism and Travel Reviews: Sri Lankan Destinations", Mendeley Data, V1, doi: 10.17632/2nbvx5m4hs.1

Data Mining Based Study to Analyze Passenger Transportation in Uva Province

W.B.M.S.C. Wijayakoon
Department of Information and Communication Technology
Uva Wellassa University of Sri Lanka
Badulla, Sri Lanka
supunw@uwu.ac.lk

W.M.C.J.T. Kuthulwatta
Department of Information and Communication Technology
Uva Wellassa University of Sri Lanka
Badulla, Sri Lanka
chirantha@uwu.ac.lk

J.K.H. Samapth
Department of Biosystems Technology
Uva Wellassa University of Sri Lanka
Badulla, Sri Lanka
hasitha@uwu.ac.lk

Kolitha B. Wijesekara
Department of Biosystems Technology
Uva Wellassa University of Sri Lanka
Badulla, Sri Lanka
kolitha@uwu.ac.lk

***Abstract*— Passenger transportation methods are a very important factor for sustainable development and sustainable supply chain management. According to the pandemic situation and to employ well established economic and social based development within the Uva Province, there is a need of identifying any data patterns, uncovered information, and obtaining other valuable information. The primary objective of this research is to uncover patterns, correlations, and trends that influence transportation operations. According to the several studies, data mining techniques were used for data analysis, pattern identification and decision making. Hence, this research analyses road passenger transportation within the Uva Province in Sri Lanka using data mining techniques. The dataset of 26,100 entries, managed by the Passenger Transport Service Authority of Uva Province from 1976 - 2022, was collected and explored to uncover patterns, correlations, and trends that influence transportation operations. This analysis reveals several significant findings. The findings contribute to evidence-based recommendations aligned with current transportation trends and market conditions. This paper underscores the significance of data mining in transforming raw data into actionable insights that facilitate informed decision-making in the transportation sector.**

***Keywords*—** *Data mining, Passenger Transportation, Analytics, Transportation patters*

I. INTRODUCTION

In light of the pandemic and the need for effective economic and social development in Uva Province, it's essential to identify data patterns and uncover valuable insights. The main goal of this research is to reveal the patterns, correlations, and trends that impact transportation operations.

Data mining has been called exploratory data analysis, among other things. Masses of data generated from cash registers, from scanning, from topic-specific databases throughout the company, are explored, analysed, reduced, and reused. Searches are performed across different models proposed for predicting sales, marketing response, and profit. Classical statistical approaches are fundamental to data mining.

Risk management, data security planning, and many other crucial business use cases all greatly benefit from data mining. It is crucial in a variety of other professions as well, including governance, sports, mathematics, and scientific research. Patterns, correlations, and trends will be found in this study's data collection on the transportation services under the control of the Passenger Transport Service Authority of Uva Province (PTSAUP), and recommendations will be made in accordance with forecasts based on transportation trends, behaviours, and market conditions discovered using data mining techniques.

All private bus transit is governed and observed by the Road PTSAUP (RPTAUP). The Ministry of Sports, Youth Affairs, Tourism, Transport, Cultural Affairs, Textiles, and Small Industries is in charge of the RPTSAUP. According to Sections No. 01 of 2001 and No. 02 of 2019 (revised) of the Statute of the Passenger Transport Service Authority, RPTSAUP was created in the year 2000. The goal was to efficiently provide passenger transport services. The RPTSAUP continues to provide services for the administrative districts of Badulla and Monaragala and maintains positive working relationships with the national Ministry of Transport and all other relevant public and private institutions in order to carry out the government's policies for the development of the transportation sector. To meet the public transportation needs of the Uva Province, 12 and 14 bus stations available in Monaragala and Badulla district respectively where 26 in total. Main research objectives of this research are as follows:

- RO1 - To discover any structure in the unstructured data set based on RPTSAUP.
- RO2 - To extract any meaning from noisy data based on RPTSAUP.
- RO3 - To discover any patterns in apparently random data on the data set.
- RO4 - To understand any patterns, correlations, and trends on the selected data set.
- RO5 - To predict any transportation trends, behaviours, and market states.

II. Literature Review

The practice of drawing insightful conclusions from huge databases is known as data mining. It has been widely used in a variety of industries, including transportation, where it is used to examine trends and behaviours in the industry.

To find a solution in the real world of transportation, numerous pieces of information must be combined. Decision-makers for transportation engineers now have additional options thanks to recent studies in transportation data mining. K-means clustering [1], decision trees [2], and neural networks [3] were used to forecast freeway trip time with non-recurrent congestion [4].

Data mining techniques such as clustering and classification were utilized in public transportation data to analyse passenger travel behaviour [5]. Similar study used clustering analysis to group passengers based on travel patterns and identified key factors affecting passenger travel behaviour. The study found that data mining could be an effective tool for bettering public transportation services and comprehending passenger behaviour [6].

In a different study, researchers employed clustering techniques to discover a relationship between the route taken by travellers and the flow of passengers at the station [7]. Furthermore, in [8] authors collected data in a metro system to estimate passenger route choice patters where probabilistic model was developed to estimate from empirical analysis. Six phases proposed in [9,10] that can be taken into account for data mining initiatives, according to the Cross-Industry Standard Procedure for data mining is categorized as, Business Understanding, Data Understanding, Data Preparation, Modelling, Evaluation, and Deployment as shown in Fig. 1.

In the field of transportation engineering, a lot of information has been gathered about traffic, accident analysis, asphalt conditions, road durability and stair management. It is identified that bus optimization and predict number of passengers in [11] used MARS algorithm to predict optimal demand. Random Forest model used in [12] to better predict regression. Furthermore, Naïve bayes model to forecast bus demand [13] and LSTM forecasting model [14] to predict short-term public transportation demand.

In conclusion, it has been demonstrated that data mining is an effective technique for examining trends and behaviours in transportation data. This study highlights the need for research in this area and offers additional proof of the utility of data mining in the transportation sector.

III. Methodology

Our stated research methodology is to find

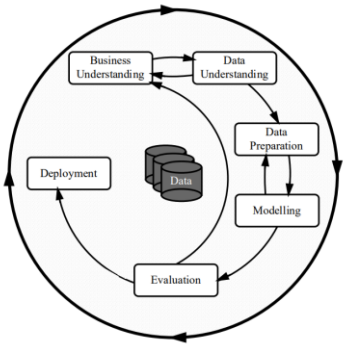


Fig. 1. Data mining model [10]

structure in chaotic data, from noisy data to find patterns and connections, and predict trends and behaviors. Mainly focused on the most recent data mining techniques put forth by researchers and consult

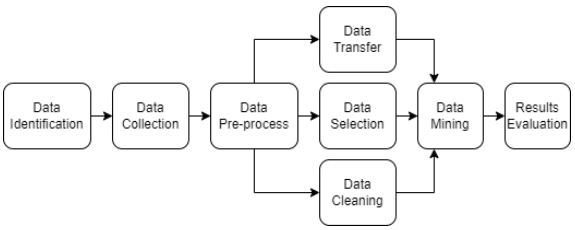


Fig. 2. Methodology

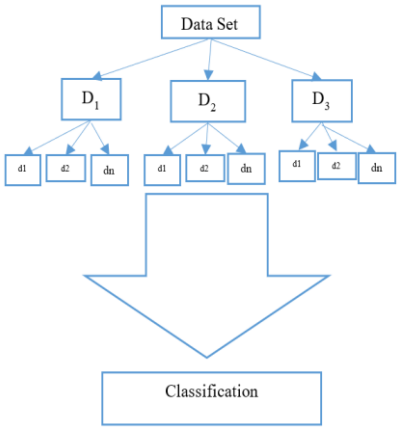


Fig. 3. Data classification steps

recently published research articles to gain the necessary knowledge to create well-organized research. As the main research steps in the methodology as shown in Fig. 2, data from the

RPTSAUP was identified, collected and preprocessed. Then applied data classification and clustering techniques. Furthermore, ZeroR classifier was used to visualize patterns and trends available in data.

A. Data Collection, Pre-Processing and Data Mining

Initially, a dataset of 26,100 entries was collected from the RPTSAUP at Uva Provincial Council and translated from Sinhala language to English language. Next, null values were cleaned, followed by the classification of the data. Data classification steps were illustrated in Fig. 3. These data are mainly classified into three parts as, bus details (8100 data), bus owner details (9000), and bus driver and conductor details (9000). Using the number of bus seats in the Bus Details data set, buses were divided into three types. These types are one-door buses, two-door short buses, and two-door long buses. They were divided into Badulla and Monaragala districts according to the locations of bus stands and divisional secretariats. The bus driver's and conductor's ID numbers were used to determine their age limits.

IV. RESULTS AND DISCUSSION

A. Bus Details Based on Year of Manufacture

The bus fleet of PTSAUP bus operations consists of over 1016 buses from different manufacturing years which range between 1976 to 2022.

According to Fig. 4, highest number of the buses in Uva Province were manufactured in 2011, which is 132. Majority of the buses were manufactured between 2000 and 2018. There are very few buses used which manufacture year is 2019 onwards.

B. Analysis of Bus Details based of the Number of Seats and Bus type:

PTSAUP Bus Operations bus fleet comprises over 1016 buses, from different manufacturers. All buses are 102 in. (2.59 m) wide. Most of these bus models consist of double seats on each side. The driver's seat is isolated and the back row usually has more seats than others. The typical bus seat is 39 inches wide and generally considered to have a maximum seating capacity of three. An average bus carries between 30 and 100 passengers.

Fig. 5 shows the number of seats in buses in Uva province. Most of the buses in Uva Province have 42 seats. It is two doors (short) busses. 303 buses belong to it. There are also 248 buses with 54 seats. It is two doors (long) busses. There are fewer single-door buses as compared to other buses.

There are three types of buses in Uva Province as one-door buses, two-door short buses, and two-door long buses. The most common type is the two doors (short) bus. A conventional door, also known as a regular door is a type of door that is hinged at the front-facing edge of the door, and so

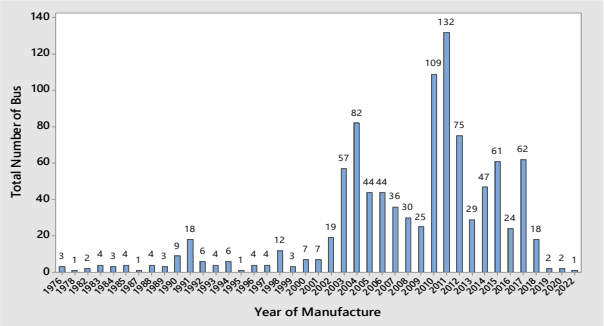


Fig. 4. Representation of buses by year of manufacture

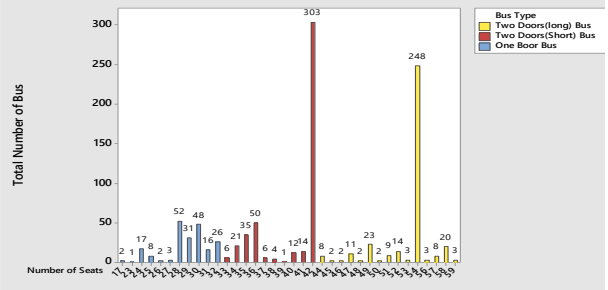


Fig. 5. Number of bus count based on bus type and number of seats available

allows the door to swing outward from the hull of the bus. These doors are relatively safe, in that if they are opened during forward motion of the vehicle, the wind resistance will work against the opening door, and will effectively force its closure. The purpose of having two doors on a bus is that it saves time at stops if boarding passengers don't have to wait for those alighting to get off.

Fig. 6 shows the type of buses. This graph clearly shows the greatest number of buses is two doors (short) buses.452 buses belong to it. There are also 358 buses with two doors (long buses). There are also 206 buses with one door in Uva province. In general, bus types are distributed as 20% are one-door buses, 45% are short buses, and 35% are long buses.

C. Analysis of Bus Details based on the Provided Service:

Buses were clustered in to 6 based on the service that they provide. They are Air-Conditioned Buses (AC), Normal Service (N), Normal and Sisuseriya (N/S), Waiting Buses (P), Nisiseriya Service (Nisi S), Sisuseriya (S).

Fig. 7 shows the clustering of bus data based on the service provided by the bus. This clearly shows that the majority of them were normal service providers which is 88%. There were 889 normal service providers, 61 Sisuseriya buses, 30 AC buses and 29 Normal /Sisuseriya buses in Uva Province. However, there were 6 Waiting buses in uva province and one bus to Nisiseriya service.

D. Analysis of Data Based on Bus Stands

A bus stand is typically used to allow a bus to stop at a bus stop. Additionally, bus stops permit

short-term parking for driver changes or breaks. A bus stand is the point where a bus route starts or ends, where they stop, turn or reverse, and wait before departing on their return journeys. It's also a place for passengers to board and alight from vehicles. It also often provides a convenient point where services can be controlled from.

Fig. 8 shows the bus stops of existing buses in Uva Province. At present, transport service facilities are being provided

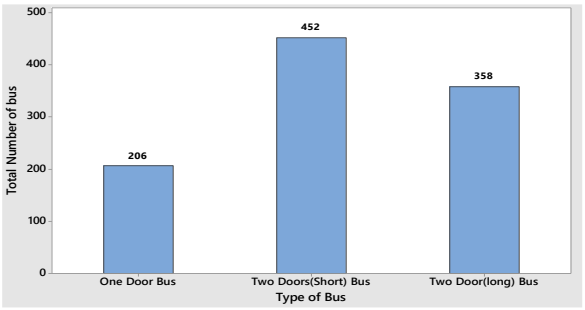


Fig. 6. Total number of bus count based on bus type

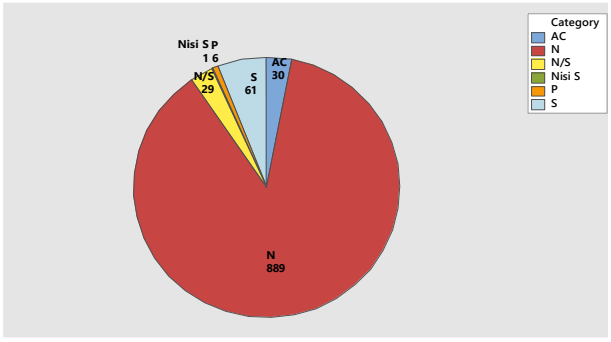


Fig. 7. Number of bus count based on the providing service

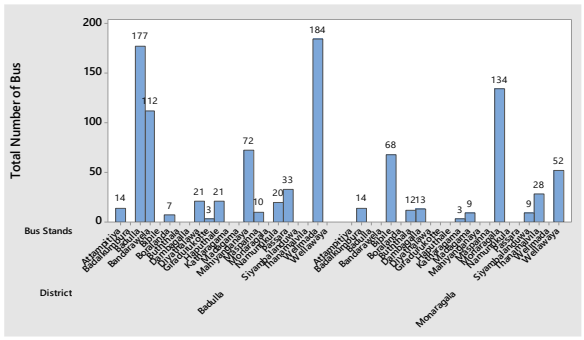
through 26 bus stations in both Badulla and Monaragala districts to fulfil the passenger transport public service needs of Uva Province. Welimada bus station has 184 buses, making it the bus station with the most buses. Next to that, Badulla, Monaragala, Bandarawela, Mahiyanganaya, and Bibile bus stations contains the most buses. The number of buses belonging to them are 177, 134, 112, 72, and 68, respectively. Out of 26 bus stations in Uva province, buses were registered for only 22 bus stations.

E. Divisional Secretariat Division (DSD)

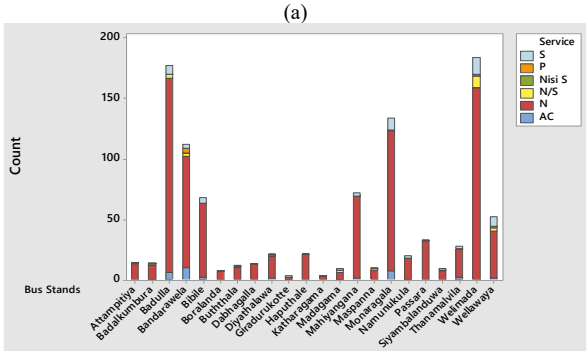
Divisional secretariats are the administrative subdivisions that make up Sri Lanka's districts. These were initially based on the *korales* and *ratas*, two types of feudal counties. Previously, they were referred to as "D.R.O. Divisions" in honour of the "Divisional Revenue Officer". Later, the D.R.O.s changed their title to "Assistant Government Agents", and the Divisions adopted the name "A.G.A. Divisions". The divisions are currently managed by a "Divisional Secretary" and are referred to as "D.S. Divisions". In Uva Province, there were 27 DSDs identified by district. Among them, 11 in the Monaragala district and 16 in the Badulla district.

Based on Fig. 9, DSD of Welimada has the highest number of buses which is 103 and Lunugala DSD has the least number of buses which is 3.

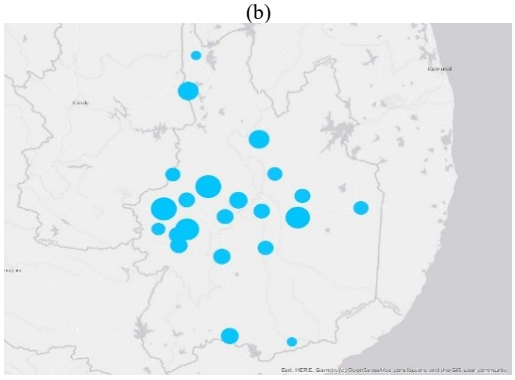
According to the Fig. 10, highest number of bus count registered under Badulla district holding 66% from total number of buses. Out of 1016 buses registered under the PTSAUP, 674 buses belong to the Badulla district, and 342 buses belong to the Monaragala district.



(a)

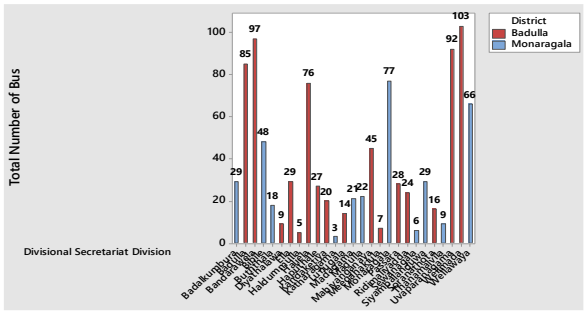


(b)

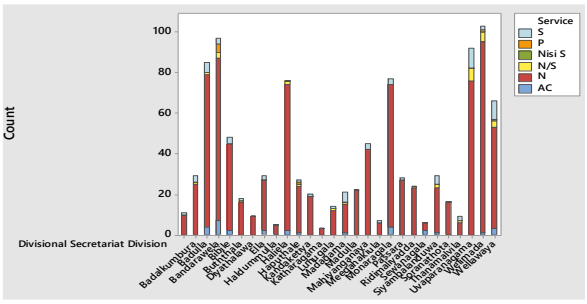


(c)

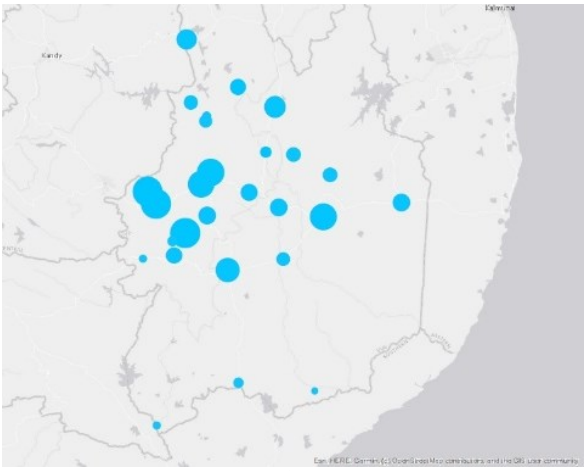
Fig. 8. (a) total number of buses registered for bus stands, (b) total number of buses registered for bus stands with their service. (c) mapping of bus stands with the number of buses



(a)



(b)



(c)

Fig. 9. (a) total number of buses registered under DSDs and relevant district, (b) total number of buses registered under DSDs and service, (c) mapping of bus DSDs with the bus count

F. *Bus Routes and Distance Traveled Per Day*

Fig. 12 (a) shows the routes travelled by buses belonging to PTSAUP. Thickness of the route is generated based on the number of buses travelling along the route. Fig. 12 (b) shows the number of times (trips) buses travel per day. Maximum number of buses travel two trips per day. Furthermore, there is a bus that travels 12 trips per day which is the highest. Details of that bus is shown in Fig. 11. It is a normal service bus, and its route is Halieala to Badulla and this bus belongs to the Halilela DSD.

G. *Analysis of Bus Owner Details*

Among all bus owners majority possess one bus. Very few owners possess more than one bus. However, maximum number of buses that owns by one person is 5. Out of those five, four were served as Sisuseriya and one served as Normal Service. Fig. 13 represents the information of those 5 buses.

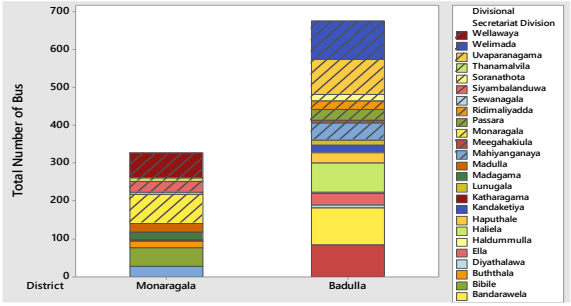


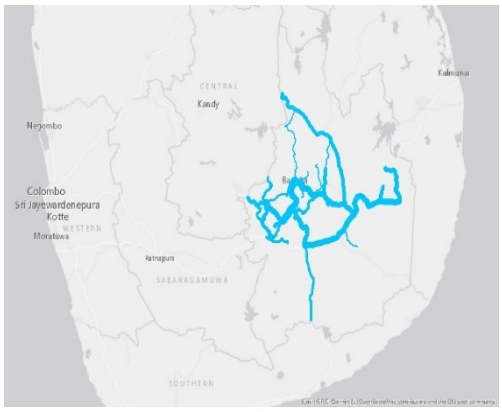
Fig. 10. Representation of the relationship between number of buses and bus stands registered under two districts

H. *Place of departure and destination*

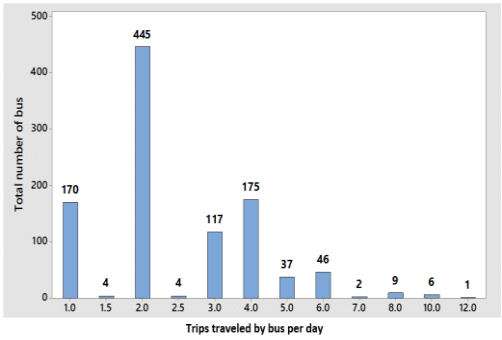
Place of departure is the point where a bus route starts and where wait before departing on their return journeys. It's also where passengers board buses. A bus place of destination is the point where a bus route ends, where buses stop or turn. It's also where passengers alight from buses. Badulla and Bandarawela are the departure and destination points for most buses in the PTSAUP. The majority of buses departure from Bandarawela and end the trip at Badulla as shown in Fig. 14 and Fig. 15.

I. Bus Drivers and Conductors Details

1) Bus Drivers: Bus drivers transport people between diverse locations, such as work, school, and towns.



(a)



(b)

Fig. 11. (a) Route mapping based on the number of buses travelled, (b) Distance travelled by bus per day

Maximum trips traveled by bus per day = 12

Divisional Secretariat Division	Haliela
Road No	296
Distance (Km)	6.40
Trips traveled by bus per day	12
Route approved to run	Haliela - Badulla
Service	N
Bus Stands	Badulla

Fig. 12. Details of the bus travelled maximum trips per day

They follow a daily schedule while transporting people on regular routes along town or suburban streets. They stop frequently, often every few blocks and when a passenger requests a stop.

Fig. 16 shows the number of drivers registered in PTSAUP based on age ranges. There were 337 drivers in the age range between 34 and 43 years where it holds the highest group of drivers while 13 drivers fall in the age range from 64–73 years where it holds the lowest group of drivers.

2) Bus Conductors: Bus conductors have a range of responsibilities, which include collecting fares, suing tickets, helping passengers with baggage, lvising on destinations and ensuring the safety of assengers.

Fig. 17 shows the number of bus conductors gistered in PTSAUP based on age ranges. There ere 265 conductors in the age range of 24–33 and it the largest group of conductors among all age nges. There were 8 conductors in the age range of 4–73 and that is the smallest group of conductors.

V. CONCLUSION

Based on the results from the analysis of passenger ansportation data in Uva Province, offers valuable sights and suggestions for improving the efficiency id effectiveness of bus operations. The majority of ises in service were found to be relatively modern hich were manufactured between 2000 and 2018. It better to continue bus upgrade efforts to enhance the ervice quality and passenger comfort. eographically, Badulla district becomes the hub for us operations while having the highest oncentrations of buses, especially in the Welimada

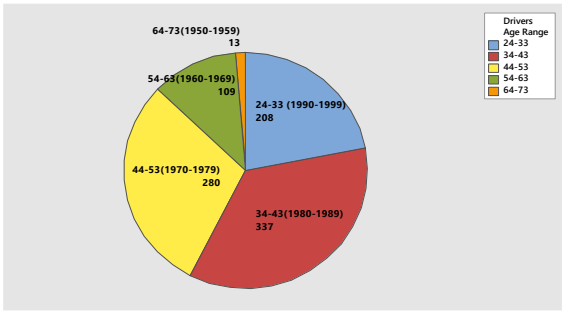


Fig. 16. The number of drivers based on age range

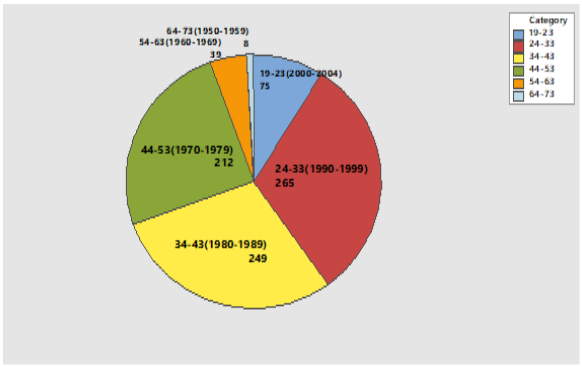


Fig. 17. The number of conductors based on age range

sector in Uva Province to improve public transportation services and enhance the overall transportation system.

The application of data mining techniques was facilitated by Minitab and Weka software to enable robust analysis. Furthermore, ArcGIS software was used for bus route mapping from patterns obtained by data clustering and operational patterns by demonstrating the potential of advanced analytics in enhancing public transportation management and service delivery.

ACKNOWLEDGMENT

We would like to express our gratitude to the Uva Wellassa University of Sri Lanka for the financial support of this research (UWU/RG/2022/19) and the Road Passenger Transportation Authority of Uva Province for providing required data and information.

REFERENCES

[1] R. A. Kadir, Y. Shima, R. Sulaiman, and F. Ali, "Clustering of public transport operation using K-means," in 2018 IEEE 3rd International Conference on Big Data Analysis (ICBDA), 2018, pp. 427-432.

[2] J. de Oña, R. de Oña, and G. López, "Transit service quality analysis using cluster analysis and decision trees: a step forward to personalized marketing in public transportation," Transportation, vol. 43, pp. 725-747, 2016/09/01 2016.

[3] D. Shmueli, "Applications of neural networks in transportation planning," Progress in Planning, vol. 50, pp. 141-204, 1998/10/01/ 1998.

[4] C.-S. Li and M.-C. Chen, "A data mining based approach for travel time prediction in freeway with non-recurrent congestion," Neurocomputing, vol. 133, pp. 74-83, 2014/06/10/ 2014.

[5] Z. Hu, L. Jingen, and H. Bing, "Analysis of passenger travel behavior based on public transportation OD data," presented at the Proceedings of the 6th International Conference on Information and Education Technology, Osaka, Japan, 2018.

[6] X. Cao, H. Chen, F. Liang, and W. Wang. (2018, Measurement and Spatial Differentiation Characteristics of Transit Equity: A Case Study of Guangzhou, China. Sustainability 10(4).

[7] N. Lathia, J. Froehlich, and L. Capra, "Mining Public Transport Usage for Personalised Intelligent Transport Systems," in 2010 IEEE International Conference on Data Mining, 2010, pp. 887-892.

[8] J. Zhao, F. Zhang, L. Tu, C. Xu, D. Shen, C. Tian, et al., "Estimation of Passenger Route Choice Pattern Using Smart Card Data for Complex Metro Systems," IEEE Transactions on Intelligent Transportation Systems, vol. 18, pp. 790-801, 2017.

[9] G. S. Linoff and M. J. A. Berry, Data Mining Techniques. For Marketing, Sales and Customer Support: Wiley Computer Publishing, 1997.

[10] R. Wirth and J. Hipp, "CRISP-DM: Towards a Standard Process Model for Data Mining," 2000.

[11] M. F. Khan, S. Asghar, M. I. Tamimi, and M. A. Noor, "Multi-Objective Transport System Based on Regression Analysis and Genetic Algorithm Using Transport Data," IEEE Access, vol. 7, pp. 81121-81131, 2019.

[12] C. Xue, Y. Ju, S. Li, and Q. Zhou. (2020, Research on the Sustainable Development of Urban Housing Price Based on Transport Accessibility: A Case Study of Xi'an, China. Sustainability 12(4).

[13] T. Stadler, A. Sarkar, and J. Dünnweber, "Bus Demand Forecasting for Rural Areas Using XGBoost and Random Forest Algorithm," in Computer Information Systems and Industrial Management, Cham, 2021, pp. 442-453.

[14] J. D. Caicedo, M. C. González, and J. L. Walker, "Public Transit Demand Prediction During Highly Dynamic Conditions: A Meta-Analysis of State-of-the-Art Models and Open-Source Benchmarking Infrastructure," arXiv preprint arXiv:2306.06194, 2023.

Ethical Landscape of Artificial Intelligence: A Review

Dinoo Gunasekera
Department of Informaiton and Communication Technology, Faculty of Technology
University of Ruhuna
Matara, Sri Lanka
dinoog@ictec.ruh.ac.lk 0000-0001-7007-6957

***Abstract*— Artificial intelligence (AI) offers great potential to enhance and improv many aspects of human life and society. However, to ensure its benefits are realized in a lawful, fair, and trustworthy manner, the ethical challenges arising from increasingly autonomous technologies must be addressed. This paper reviews the current state of AI ethics research and ongoing multidisciplinary discussions. Author identified obstacles inhibiting responsible progress and opportunities to foster innovation aligned with principles like transparency, fairness, privacy, and human rights. The literature review examined technical reports and peer-reviewed research on issues such as biases in training data that could lead to unfair outcomes, challenges in ensuring accountability and oversight of advanced autonomous systems, and the tensions between maximizing social benefits and mitigating risks to privacy, fairness, and other human rights. The author found that these are key obstacles to achieving trustworthy AI applications. Emphasizing data and algorithmic transparency, inclusive governance through collaboration in technical and policy solutions, and user-centered approaches aligned with societal values offers a path towards safely and equitably realizing AI's promise. This can be achieved through flexible yet robust frameworks for managing risks in dynamic human-machine systems.**

Keywords*— *Artificial Intelligence, Ethics, Data Bias in AI, Data Privacy

I. INTRODUCTION

In the rapidly evolving landscape of artificial intelligence (AI), where innovation intersects with ethical considerations, a thorough exploration of the ethical dimensions becomes imperative. As AI technologies permeate diverse facets of our lives, from healthcare to finance, education to law enforcement, the ethical implications of their development and application take center stage. The development of AI has brought up difficult ethical dilemmas. Understanding the ethical consequences of AI systems' creation and use has become crucial as they impact every aspect of our society. This research embarks on a journey towards a deeper understanding of the Ethics of AI, offering a critical review that synthesizes the current state of knowledge and literature in this pivotal field.

By providing an in-depth analysis of the most recent findings and research in this crucial area, this paper aims to contribute to the ongoing conversation about the ethics of AI. The ethical challenges surrounding AI are complex and cover a wide range of topics, from responsibility and transparency to privacy and bias. This review aims to illuminate the ethical difficulties and opportunities that arise in the age of AI.

With the widespread adoption of AI technologies, addressing the ethical implications of their development and use has become increasingly crucial [4]. Failure to do so could lead to gaps in transparency, safety, and ethical standards. These concerns have prompted various initiatives and guidelines, such as the Ethical Guidelines for Trustworthy AI by the European Commission and Microsoft's FATE: Fairness, Accountability, Transparency, and Ethics in AI. Trustworthy AI is based on seven technical requirements sustained over three main pillars that should be met throughout the system's entire life cycle: it should be lawful, ethical, and robust, both from a technical and a social perspective [9].

Furthermore, the impact of AI on society, such as changes in the job market and potential misuse by malicious actors, has also raised ethical concerns. Addressing these ethical implications requires collaboration among developers, policymakers, and society at large. The field of AI ethics is increasingly crucial as we tackle the manifold potential harms society has faced due to AI systems-and as we work to pre-empt future harms [28]. To ensure the responsible and ethical development and use of AI systems, it is necessary to establish ethical guidelines and regulatory frameworks that promote fairness, transparency, and accountability [5,27]. These efforts should also include engaging in critical discussions about the societal impact of AI and empowering individuals to actively participate in shaping the future of AI in a responsible and ethical manner.

Several key organizations and initiatives have emerged to address the ethics of AI. These include the IEEE initiative for ethically aligned design of autonomous intelligent systems, the European Commission's Ethics Guidelines for Trustworthy AI, and the World Health Organization's exploration of ethical challenges in AI for healthcare [11]. Leading tech companies such as IBM, Google, and Microsoft

have also published frameworks and principles to guide the deployment of AI systems and have established research positions and units focused on AI ethics [23].

For example, the Institute for Electrical and Electronics Engineers established the Global Initiative on Ethics of Autonomous and Intelligent Systems in 2016. This initiative aims to promote the ethical development and deployment of autonomous and intelligent systems. Various governmental organizations, associations, and private companies have proposed guidelines and codes of ethics for the design and use of AI. These efforts reflect a growing awareness of the need for ethical considerations in AI and the recognition that responsible development and deployment of AI systems require guidelines, regulations, and collaboration between different stakeholders.

A vast body of cross-disciplinary research over the past decade has significantly furthered understanding of AI ethics' technical and societal dimensions. Among the earliest influential works were analyzed of harmful biases found within popular word embedding techniques still broadly utilized today [17]. By empirically demonstrating how inherent prejudices from language data become encoded within these widely used AI building blocks, discussions on mitigating unfair impacts from model training were elevated. Follow-up studies introduced techniques seeking to remedy gender stereotypes within the embeddings [6], but more recent analyses note biases remain an evolving challenge [13].

Pioneering examinations of membership inference attacks on language models illuminated privacy issues requiring amends as well [15]. As AI infiltrates new domains involving sensitive personal data, from biometrics to conversational dialog, ensuring strong consent and control over information use stands paramount [7]. Fairness considerations extend beyond mitigating biases within models, as predictive tools risk exacerbating real-world inequities if deployed without oversight [21]. Technical strategies for auditing subgroup impacts are emerging but promising [2,19].

Cross-cutting themes of transparency, accountability, and meaningful participation in governance have crystallized as core priorities moving forward. While challenges abound, balancing competing objectives demands open-minded, multistakeholder cooperation. Reviews flag a lack of standard practices and metrics for evaluating progress [18]. Some advocate "build fast but build safe" through techniques like value-aligned design and constitutional AI to better reason about societal impacts [26]. International policy initiatives aim to foster responsible innovation through Human-Centered AI frameworks emphasizing domains like healthcare, education and justice [8].

This research considers technical, policy and social dimensions of bias, privacy, fairness, transparency and long-term effects - both obstacles and promising solutions. By comprehensively analyzing debates and mapping the landscape, this paper explores how diverse perspectives might collectively strengthen safeguards to help maximize AI's benefits for all.

In conclusion, the necessity for an in-depth understanding of AI's ethical implications has never been more important as its influence multiplies. To further inform moral AI development and use, this study aims to investigate the ethical issues surrounding AI, present a state-of-the-art perspective on AI ethics research, and analyze the ethical considerations surrounding AI.

II. Foundation of Ethics in Artificial Intelligence

Ethics refers to moral principles that guides a person's or group's behavior, aiming to distinguish between "right" and "wrong" actions to minimize harm and promote human well-being. Understanding the philosophical foundations of ethics provides important context for discussions around AI ethics as these technologies progress.

Three influential schools of ethical thought are consequentialism, deontology, and virtue ethics. Consequentialism focuses on outcomes, holding that the moral worth of an action depends on its results [29]. Deontology emphasizes duties and rights, judging actions based on adherence to rules or principles irrespective of outcomes. Virtue ethics examines moral character, focusing on virtues or character traits like honesty, mercy or fairness.

Different domains of application may draw from multiple perspectives, but a common goal is aligning technical systems with societal values like fairness, safety, privacy and transparency. As AI capabilities advance, these questions have grown in relevance and complexity, given machines' expanding roles directly impacting people's lives. Researchers in the late 2000s began emphasizing the need for "machine ethics" [22] tailored to AI's unique affordances and risks.

Over subsequent years, recognition grew that building accountable technologies demands stakeholder participation from developers, policymakers, social scientists and beyond. Discussions evolved from examining general AI safety issues to focusing on specific applications and use cases. Frameworks emerged highlighting needs like non-discrimination, user control, and transparency across contexts from healthcare to criminal justice to consumer services [12]. While challenges remain as the field moves rapidly, integrating ethical safeguards from concept to deployment aims to ensure AI is "developed and applied safely, for and with humans" [1].

Grounding ongoing conversations in the philosophical underpinnings of ethics helps strengthen conceptual rigor and progress assessment. Emphasizing empirical approaches balanced with multidisciplinary cooperation aims to resolve complex

uncertainties. Systematically addressing associated technical, policy and social considerations offers the most viable path towards realizing artificial intelligence's vast benefits safely and equitably for all.

III. Ethical Issues and Challenges in Artificial Intelligence

As artificial intelligence technologies take on more autonomous roles affecting human lives, addressing the associated ethical issues becomes crucially important. However, integrating responsible practices faces significant challenges across technical, policy and social dimensions. A lack of standard practices remains problematic as the field advances rapidly through different contexts.

A. Holistic Consideration and Societal Challenges:

Earlier approaches mainly focused on transparency and fairness but fell short of addressing the complexity of biases. Mitigation requires holistic consideration of socio-economic factors, like the power structures shaping data. Technical solutions alone cannot remedy deeper societal issues [25]. Additionally, formalizing concepts like fairness and accountability amid uncertainties in rapidly progressing AI abilities remains challenging.

B. Addressing Biases and Value Judgments:

Addressing biases involves value judgments that prioritize certain groups, necessitating inclusive governance [18]. Current guidelines differ substantially in scopes and levels of enforceability. Monitoring real-world impacts also poses difficulties, especially when attributing outcomes to human or algorithmic factors.

C. Multistakeholder Cooperation:

Proposed solutions center on multistakeholder cooperation. Frameworks evaluated over 60 guidelines and found convergence around transparency, human oversight, privacy, and fairness principles, but there is variance in implementation guidance [18]. Others detail "value sensitive design", emphasizing iterative assessment integrating technical, empirical and conceptual methods [25].

Additional recommendations include emphasizing data provenance, oversight throughout planning and execution, and accounting for global sociocultural differences [12]. International cooperation remains crucial as technologies transcend borders. Standardized self-assessment and benchmarking processes could help evaluate progress while balancing innovation.

D. Data and Representation Issues:

As discussed by Jarrahi et al. [16], inherent biases in the sourced data used to develop AI systems can encode unfair prejudices. This is a major concern as models may discriminate or disadvantage certain groups without inclusion of diverse perspectives in data collection practices. Representation imbalances can impede impartial, equitable outcomes if left unaddressed. Russell [14, 24] notes the complexity of many advanced AI systems makes singling out responsible parties difficult when things go wrong. Their autonomous, adaptive natures compound this issue as unanticipated emergent behaviors may arise. Without robust accountability frameworks, oversight becomes ambiguous, risking unpremeditated harms.

E. Reconciling Competing Priorities:

Jobin et al. [18] explore the tensions policymakers frequently face balancing objectives like privacy, fairness, public benefit, and innovation. No Silver bullet exists, and trade-offs often must be considered situationally based on impacts. The lack of universal metrics for priority-setting adds to the challenging of coherent governance amid conflicting stakeholder views.

F. Alignment Challenges:

Rapidly changing AI capabilities push the boundaries of current methods for embedding safeguards. Controls proven effective today may not transfer as technology and its applications evolve, requiring anticipatory, adaptive solutioning to maintain protection aligned with progress [8].

G. Verification Difficulties:

Lifelong monitoring poses barriers, as AI upgrades and deployed use cases diversify and scale over time. Comprehensively validating controls function as intended throughout such dynamic environments needs innovative approaches for assurance throughout system lifespans [20].

Effectively addressing complex AI ethics issues demands holistic, collaborative, and adaptive approaches. Focusing on transparency, inclusion, responsibility and human rights principles provides a coherent starting point. Continuous stakeholder participation in technical and policy solutioning will maximize benefits and trustworthiness for all.

IV. Principles That Have Emerged as Foundations for Developing Ethical AI Systems

A. Machine Ethics and Safety

An early influential work by Moor [22] established the need for explicit "machine ethics" tailored to AI's abilities and risks. As AI’s applications expanded into more safety-critical decision-making roles directly impacting people, principles like safety, transparency and accountability become paramount.

B. Common Values Frameworks

Frameworks have consolidation around common values like fairness, privacy and human welfare. For example, the "AI4People" approach by Floridi & Cowls [12] emphasizes goas such as explicability to sustainability. These frameworks aim to cultivate beneficial, trustworthy technologies through participatory governance and impact assessments.

C. Global Convergence on Principles

Jobin et al. [18] found that nearly all guidelines converge on overarching principles of transparency, fairness, and human control. However, their interpretation varies across sociocultural contexts and application areas, necessitating differentiation.

D. Value Sensitive Design

Shilton [25] proposed the "Value Sensitive Design" approach, an iterative methodology integrating technical, empirical, and conceptual methods. This approach systematically embed values like dignity, autonomy, and informed consent into systems throughout all stages of development.

E. Key Ethical Principles

Integrating public values centers around key principles addressed in the literature, including:

- Transparency: Ensuring intelligibility, oversight, and accountability for all stakeholders.
- Explainability: Allowing stakeholders to understand AI system behaviors and decisions This principle enables oversight by helping users comprehend the logic, priorities, and limitations behind model outputs [18].
- Fairness and non-discrimination: Mitigating biases and promoting impartial, inclusive outcomes.
- Privacy and data governance: Incorporating consent, purpose limitation, and security safeguards around sensitive user information.
- Human welfare: Prioritizing wellbeing and aligning AI with ethics through governance that facilitates public participation and benefit.
- Autonomy: Preserving and enhancing user freedom of choice and control and control related to AI systems and algorithms impacting their lives.
- Trustworthiness: Ensuring techniques that prove reliability and assurance that system objectives and behaviors match societal expectations.

F. Ethical Considerations

- Welfare: Floridi and Cowls [12] view human welfare as the overarching goal that should guide all other principles. This encompasses well-being, rights, values, and overall quality of life. Careful impact assessments are needed to ensure AI enhances rather than jeopardizes welfare.
- Privacy and data governance: Jarrahi et al. [16] emphasizes informed consent and clear disclosure of data practices to end-users. Strong security procedures and adherence to the original purpose of data collection are crucial to upholding individual's privacy interests.
- Fairness and non-discrimination: Jobin et al. [18] link the pursuit of impartial outcomes with mitigating systemic biases that could disadvantage certain groups. Culturally informed definitions of fairness and discrimination prioritize equity for all.
- Trustworthiness and assurance: Dilmaghani et al. [10] describes trustworthiness as the standard that should substantiate AI systems' reliability, safety, and alignment with expectations over time. Principles like transparency and accountability help enable continued assurance throughout dynamic developmental and application environments.

G. Guiding Ethical Frameworks

Different ethical frameworks can guide AI development responsibilities. Floridi and Cowls [12] discuss three prominent theories:

- Consequentialism: Explored by Moor [22], this theory points that actions are right based on maximizing positive and minimizing negative outcomes. For AI, this means designing systems that benefit humanity while avoiding harm.
- Deontology: Covered by Jobin et al. [18], this view holds that duties and moral rules determine responsibilities, even if outcomes differ. Ensuring privacy through data protection rules aligns with this perspective over maximizing profit alone.
- Virtue Ethics: Shilton [25], characteristics like trustworthiness, beneficence, and wisdom should drive work. Building AI that augments users' lives exhibits the care and competence of virtuous work.

Floridi and Cowls [12] find that these frameworks can harmonize. Where consequences and rules disagree, virtue ethics propose alternatives that respect dignity and care. Together, these perspectives help navigate technical, social, and ethical complexities to safeguard people through diligent, principled innovation.

By combining technological, empirical, and conceptual approaches with inclusive multistakeholder collaboration, these foundational ethical principles aim to maximize AI's benefits while safeguarding humanity in complex, uncertain environments.

ACKNOWLEDGMENT

The completion of this undertaking could not have been possible without the participation and assistance of many individuals whose contributions are sincerely appreciated and gratefully acknowledged. The authors would like to express their deep appreciation to all who supported and encouraged this research. We are particularly indebted to Dinoo Gunasekera for her unwavering support and motivation, which have deeply inspired us. It was a great honor to conduct this research with her guidance. Finally, we extend our gratitude to everyone who helped us achieve success in this endeavor.

REFERENCES

[1] A European approach to artificial intelligence. (2023). European Commission.

[2] Adebayo, J. A. (2016). ToolBox for diagnosing bias in predictive modeling.

[3] Agbese, M., Alanen, H. K., Antikainen, J., Halme, E., Isomaki, H., Jantunen, M., Kemell, K. K., Rousi, R., Vainio-Pekka, H., & Vakkuri, V. (2021). Governance of Ethical and Trustworthy Al Systems: Research Gaps in the ECCOLA Method. Proceedings of the IEEE International Conference on Requirements Engineering, 2021-Septe(November), 224–229. https://doi.org/10.1109/REW53955.2021.00042

[4] Almpani, S., Stefaneas, P., & Frangos, P. (2022). Argumentation-Based Logic for Ethical Decision Making. Studia Humana, 11(3–4), 46–52. https://doi.org/10.2478/sh-2022-0015

[5] Annenberg, E. F. (2023). FAIRNESS AND BIAS IN ARTIFICIAL INTELLIGENCE: A BRIEF SURVEY OF SOURCES, IMPACTS, AND MITIGATION STRATEGIES Emilio. 109–115.

[6] Bolukbasi, T., Chang, K. W., Zou, J., Saligrama, V., & Kalai, A. (2016). Man is to computer programmer as woman is to homemaker? Debiasing word embeddings. Advances in Neural Information Processing Systems, 2016, 4356–4364.

[7] Chauhan, C., & Gullapalli, R. R. (2021). Ethics of AI in Pathology: Current Paradigms and Emerging Issues. American Journal of Pathology, 191(10), 1673–1683. https://doi.org/10.1016/j.ajpath.2021.06.011

[8] Communication from the commission to the European parliament, the council, the European economic and social committee and the committee of the regions empty: Building Trust in Human-Centric Artificial Intelligence. (2012). https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:52019DC0168&rid=6

[9] Díaz-Rodríguez, N., Del Ser, J., Coeckelbergh, M., López de Prado, M., Herrera-Viedma, E., & Herrera, F. (2023). Connecting the dots in trustworthy Artificial Intelligence: From AI principles, ethics, and key requirements to responsible AI systems and regulation. Information Fusion, 99(June), 101896. https://doi.org/10.1016/j.inffus.2023.101896

[10] Dilmaghani, S., Brust, M. R., Danoy, G., Cassagnes, N., Pecero, J., & Bouvry, P. (2019). Privacy and Security of Big Data in AI Systems: A Research and Standards Perspective.

[11] Dyoub, A., Costantini, S., & Lisi, F. A. (2022). Learning Domain Ethical Principles from Interactions with Users. Digital Society, 1(3), 1–12. https://doi.org/10.1007/s44206-022-00026-y

[12] Floridi, L., & Cowls, J. (2022). A unified framework of five principles for AI in society. Machine Learning and the City: Applications in Architecture and Urban Design, 1, 535–545. https://doi.org/10.1002/9781119815075.ch45

[13] Gebru, T., Morgenstern, J., Vecchione, B., Vaughan, J. W., Wallach, H., Iii, H. D., & Crawford, K. (2021). Datasheets for datasets. Communications of the ACM, 64(12), 86–92. https://doi.org/10.1145/3458723

[14] Groumpos, P. P. (2022). Ethical AI and Global Cultural Coherence: Issues and Challenges. IFAC-PapersOnLine, 55(39), 358–363. https://doi.org/10.1016/j.ifacol.2022.12.052

[15] Hilprecht, B., Härterich, M., & Bernau, D. (2019). Monte Carlo and Reconstruction Membership Inference Attacks against Generative Models. Proceedings on Privacy Enhancing Technologies, 2019(4), 232–249. https://doi.org/10.2478/popets-2019-0067

[16] Jarrahi, M. H., Lutz, C., Boyd, K., Oesterlund, C., & Willis, M. (2023). Artificial intelligence in the work context. Journal of the Association for Information Science and Technology, 74(3), 303–310. https://doi.org/10.1002/asi.24730

[17] Jentzsch, S., Rothkopf, C., Schramowski, P., & Kersting, K. (2019). Semantics derived automatically from language corpora contain human-like moral choices. AIES 2019 - Proceedings of the 2019 AAAI/ACM Conference on AI, Ethics, and Society, October, 37–44. https://doi.org/10.1145/3306618.3314267

[18] Jobin, A., Ienca, M., & Vayena, E. (2019). The global landscape of AI ethics guidelines. Nature Machine Intelligence, 1(9), 389–399. https://doi.org/10.1038/s42256-019-0088-2

[19] Kamińska, J. (2022). The impact of data bias on your business & the benefits of fair AI.

[20] Liu, A., Swayamdipta, S., Smith, N. A., & Choi, Y. (2022). WANLI: Worker and AI Collaboration for Natural Language Inference Dataset Creation. Findings of the Association for Computational Linguistics: EMNLP 2022, 6855–6876. https://doi.org/10.18653/v1/2022.findings-emnlp.508

[21] Lum, K., & Isaac, W. (2016). To Predict and Serve?

[22] Moor, J. H. (2006). The Nature, Importance, and Difficulty of Machine Ethics. 21(4), 18–21.

[23] Pei, Y. Z., & En, H. C. (2011). Services oriented approach for distributed heterogeneous information integration. Advanced Materials Research, 171–172, 715–718. https://doi.org/10.4028/www.scientific.net/AMR.171-172.715

[24] Russell, S. (2021). Artificial Intelligence and the Problem of Control. Perspectives on Digital Humanism, pp.19–24. doi:https://doi.org/10.1007/978-3-030-86144-5_3.

[25] Shilton, K. (2018). Values and Ethics in Human-Computer Interaction.

[26] Soares, N., & Fallenstein, B. (2015). Toward Idealized Decision Theory. http://arxiv.org/abs/1507.01986

[27] Stahl, B. C., Brooks, L., Hatzakis, T., Santiago, N., & Wright, D. (2023). Exploring ethics and human rights in artificial intelligence – A Delphi study. Technological Forecasting and Social Change, 191(September 2021), 122502. https://doi.org/10.1016/j.techfore.2023.122502

[28] Stix, C. (2022). Artificial intelligence by any other name: a brief history of the conceptualization of "trustworthy artificial intelligence." Discover Artificial Intelligence, 2(1). https://doi.org/10.1007/s44163-022-00041-5

[29] Tonkens, R. (2009). A Challenge for Machine Ethics.

Predictive Modeling of Tourist Arrivals in Sri Lanka Using Linear Regression

Nalinda Somasiri

Dept. of Computer Science & Data Science
York St John University London Campus
London, United Kingdom
n.somasiri@yorksj.ac.uk

Lakmali Karunarathne

Dept. of Computer Science & Data Science
York St John University London Campus
London, United Kingdom
l.karunarathne@yorksj.ac.uk

Swathi Ganesan

Dept. of Computer Science & Data Science
York St John University London Campus
London, United Kingdom
s.ganesam@yorksj.ac.uk

***Abstract* - One of the sectors of a developing nation's economy that directly generates income is tourism. For this reason, predicting the number of visitors is crucial when deciding on policies to upgrade facilities and other relevant aspects of this sector. The data for this paper were gathered from the Corporate website of the Sri Lanka tourism development authority(SLTDA)This paper attempts to forecast tourist arrivals in Sri Lanka using Linear Regression Model. The time span used for this study is from January 2021 to March 2024. The performance of the model was evaluated using metrics such as Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE) and R2 score. This paper forecasts tourist arrivals in Sri Lanka using the Linear Regression Model. While Linear Regression is suitable due to its simplicity and the scope of the dataset, more advanced forecasting models such as ARIMA or deep learning approaches like LSTM could have been considered for a more robust comparison. The study uses data from January 2021 to March 2024, and the model's performance is evaluated using metrics such as Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and R^2 score. The findings suggest that machine learning models can improve the forecasting accuracy of visitor arrivals, providing valuable insights for stakeholders in the tourism industry.**

Keywords: Linear Regression Model; forecast; tourist arrivals; predictions; MAE; RMSE; R2 score

I. INTRODUCTION

Sri Lanka's economy, like many others across the world, depends heavily on the tourism industry. Sri Lanka's tourist sector makes a substantial contribution to the country's GDP, employment, and foreign exchange profits. The island's distinctive wildlife, varied landscapes, and rich cultural legacy draw millions of visitors each year, making it one of the most dynamic areas of the national economy. Precise forecasting of visitor numbers is crucial for several reasons. Such as economic planning, resource management, policy formulation, crisis management and more. Better economic forecasting and planning are made possible by the ability to predict visitor arrivals. It ensures that companies, the government, and infrastructure may adjust resources, infrastructure, and services in response to variations in the number of tourists. Stakeholders can more efficiently manage resources, such as lodging, transportation, and attractions, by knowing predicted visitor inflows. This contributes to upholding excellent service standards and improving the traveler experience. Precise forecasts aid authorities in developing plans to boost travel, upgrade infrastructure, and guarantee environmentally friendly tourism practices. It also aids in the creation of regulations to lessen the negative impacts of tourism, such as deterioration of the environment and dilution of culture. The travel and tourist industry are susceptible to a range of external shocks, including natural disasters, pandemics, and geopolitical crises. By offering early warnings and assisting in the creation of backup plans, predictive modelling can help with crisis management [1], [2], [3].

Predictive modelling using machine learning approaches has demonstrated encouraging results in a number of sectors, including tourism, in recent years. Compared to conventional statistical techniques, machine learning models provide more reliable and accurate predictions by analyzing intricate patterns and correlations inside huge datasets.

Using linear regression, one of the foundational and extensively used machine learning techniques, this study attempts to create a predictive model for tourist arrivals in Sri Lanka. The model endeavors to furnish dependable forecasts that may facilitate strategic deliberation and planning in the tourist domain by using past data and diverse economic factors.

A. Research Objectives

The main goal of this research is to use machine learning techniques, notably linear regression, to create a reliable predictive model for the number of tourists that will arrive in Sri Lanka. The goal of the project is to produce accurate and trustworthy estimates of tourist arrivals by examining historical data on arrivals as well as a range of economic factors. This will enable stakeholders to plan ahead and anticipate changes in the demand for tourism.

Furthermore, in order to provide important insights into the dynamics of the tourism industry, the study attempts to identify and comprehend the major external and economic factors impacting the number of tourists that arrive. Based on the insights obtained from the model, this predictive capability is meant to assist enterprises, tourism authorities, and policymakers in making well-informed decisions about infrastructure development, resource allocation, and strategic planning. Additionally, by giving early warning signals through precise predictions, the study seeks to enhance stakeholders' capacity to manage and reduce the effects of external shocks, such as pandemics and economic downturns. By enabling improved planning and management of tourist intakes, this study ultimately hopes to contribute to the development of sustainable tourism practices by striking a balance between economic rewards and environmental and cultural preservation. By taking an all-encompassing approach, we hope to show how machine learning can be used to forecast tourism trends and promote the sustainable growth and development of Sri Lanka's tourism sector [1], [2].

II. LITERATURE REVIEW

For many nations, like Sri Lanka, where it contributes significantly to GDP, employment, and foreign exchange profits, tourism is an important economic industry. Predicting visitor numbers with precision is essential for economic planning, resource management, and policy development. This study of the literature looks at the situation of tourism in Sri Lanka right now, looks at earlier research on visitor arrivals, and emphasizes the economic factors that affect tourism patterns.

A. Tourism Industry in Sri Lanka

Over the past few decades, Sri Lanka's tourism industry has grown significantly, propelled by the island's unique wildlife, varied landscapes, and rich cultural legacy. The Sri Lanka Tourism Development Authority (SLTDA) reports that there has been a steady growth in the number of tourists arriving, with particular surges in the years leading up to the COVID-19 pandemic. The SLTDA offers thorough data on a range of tourism-related topics, such as visitor demographics, spending trends, and seasonal variations, all of which are essential for comprehending the workings of the sector [1].

The Central Bank of Sri Lanka (CBSL) offers economic indicators that shed more light on the financial aspects of the travel and tourist industry. Important indicators like GDP share, inflation rates, and currency rates have a direct impact on travel-related activities. The economic indicators of the CBSL show how closely related the tourism industry is to other economic sectors, underscoring the necessity of a comprehensive approach to tourism management and policy-making [2].

B. Economic Indicators Influencing Tourism

Arrivals of tourists in Sri Lanka are influenced by a number of economic factors, such as monetary policy, GDP growth, inflation, and exchange rates. The World Bank offers a wealth of information about Sri Lanka's economic circumstances, including insights into long-term patterns and structural shifts in the country's economy. For example, the affordability and appeal of Sri Lanka as a travel destination can be impacted by GDP growth rates and inflation levels. Changes in exchange rates can affect how affordable the resort is for tourists from other countries [3].

Understanding the macroeconomic conditions affecting tourism requires knowledge of the money supply, interest rates, and fiscal policies, all of which are covered in detail in the Central Bank of Sri Lanka's economic indicators report. A shift in the central bank rate or the average weighted fixed deposit rate may have an indirect effect on tourism by influencing investment conditions and general economic stability [2].

C. Previous Studies on Tourist Arrivals in Sri Lanka

In emerging nations, tourism has become a major engine of economic expansion, created jobs and increased foreign exchange reserves. Sri Lanka's post-war economic recovery is heavily dependent on tourism, which makes precise forecasts of foreign visitor arrivals necessary for informing business environments, infrastructure development, and successful governmental decisions.

In their paper "Forecasting international tourist arrivals in formulating tourism strategies and planning: The case of Sri Lanka," the authors create accurate forecasting models for both the top 10 source countries and total international arrivals using the Seasonal Autoregressive Integrated Moving Average (SARIMA) method. The training dataset consisted of monthly data from January 1984 to December 2016, and the accuracy of the model was tested using data from January 2017 to December 2017.
The SARIMA approach was selected because of its resilience in capturing the intricate seasonal patterns present in tourism data, as well as its reputation for handling seasonal changes. The results highlight a number of important insights such as four million tourist arrivals by 2020 is an ambitious goal set by the Sri Lankan government, forecasting accuracy is critical to strategic tourism planning and policy

formulation and SARIMA models show efficacy in producing accurate forecasts even in the face of seasonal fluctuations. By illustrating the usefulness of SARIMA models in predicting the number of foreign visitors, this study makes a substantial contribution to the literature on tourism. Governmental organizations and private sector stakeholders engaged in infrastructure planning, advertising campaigns, and policymaking in the tourism sector can benefit from the actionable insights this research offers. Stakeholders can optimize resource allocation and strategic actions to optimize the socio-economic advantages of tourism in Sri Lanka by utilizing these forecasting methodologies [4].

The study "Time Series Patterns of Tourist Arrivals to Sri Lanka" examined monthly time series data from January 2008 to December 2014, with a particular focus on the top four market regions—Asia, Western Europe, Eastern Europe, and the Middle East—for tourist arrivals. To find trends in arrivals in various regions, the study used Time Series plots, Auto-Correlation Functions (ACF), and descriptive statistics. The study's approach involved comparing the mean number of tourists arriving in each of the chosen regions using one-way Analysis of Variance (ANOVA). The results showed that, with 29,361 and 25,982 average arrivals, respectively, Asia and Western Europe had the highest arrival rates, with no discernible difference between the two regions. On the other hand, average arrivals in the Middle East and Eastern Europe were lower, at 4,300 and 5,866, respectively. Important distributional features of tourist arrivals were also emphasized by the study: arrivals from Asia, Western Europe, and Eastern Europe were non-normally distributed and showed positive skewness. Additionally, it was discovered that the data series for all four regions were non-stationary, highlighting the dynamic character of visitor flows to Sri Lanka. Also, this study suggested investigating more complex time series forecasting techniques in the future, including moving averages, exponential smoothing, decomposition methods, linear and non-linear trend models, and circular models. These techniques are suggested in order to improve forecast accuracy and identify the cyclical and seasonal trends present in Sri Lankan tourism arrivals [5].

A recent study by Basnayake et al. (2023) employs a time-series approach with Change Point Analysis (CPA) and Seasonal Autoregressive Integrated Moving Average (SARIMA) models to model and forecast tourist arrivals in Sri Lanka. The analysis makes use of monthly data on visitor arrivals from January 2000 to December 2019, a time that was remarkable for important events like the tsunami and the civil war, both of which had a major impact on tourism numbers. The goal of the authors' SARIMA model is to adequately capture seasonal fluctuations in visitor arrivals by extending standard ARIMA models to integrate seasonality. The best-performing models are chosen using the Akaike Information Criterion (AIC), and their accuracy is assessed using the Normalized Root Mean Squared Error (NRMSE) and Mean Absolute Percentage Error (MAPE) [6].

III. METHODOLOGY

Time series data of tourist arrivals, denoting visits of foreigners to Sri Lanka for the purpose of holidays, business, sight-seeing, excursions, medical treatment, religious pilgrimages, and other purposes that are uniquely different from their usual place of residence, were used to forecast future trends. The data spans from January 2021 to March 2024 , obtained in monthly time intervals from the Sri Lanka Tourism Development Authority. The secondary data for the study was collected from publications, reports, books, journals, and relevant websites that included data on tourist arrivals. Then the data was prepared for analyzing by finding null values, duplicate values, rearranging datatypes at the data pre preparation stage. After that the data set was analyzed by splitting it into training and testing sets and employed the Linear Regression model to get the forecast tourist arrivals in Sri Lanka. The Linear Regression Model was chosen for its simplicity and interpretability. More advanced models like ARIMA or LSTM, which could account for non-linearities in the data, were considered but excluded in favour of Linear Regression due to the linear nature of the dataset and the study’s focus on simplicity. However, these models offer benefits like handling seasonality and long-term dependencies, making them worthy of exploration in future work.
The dataset was pre-processed by addressing missing data using mean imputation and removing duplicates. Handling outliers was done by reviewing statistical distributions, ensuring the cleaned data is suitable for predictive modelling. Expanding the details on how the dataset was pre-processed improves the study’s reproducibility.

A. Data Collection

The study's dataset includes historical information on visitor arrivals to Sri Lanka from January 2021 to March 2024 and the dataset contains 39 rows and 13 columns. The 13 columns contain Date, Head Inflation, Core Inflation, Tourist Arrivals, Covid, GDP, %GDP, Rev(%GDP), Expenditure(%GDP), Money Supply(Annual Change), Money Supply(%GDP), Average Weighted Fixed Deposit Rate (AWFDR) Rate(%) and the Central Bank Rate. The Head Inflation columns represent the overall inflation rate in the economy and normally it measured by the consumer price index (CPI). It reflects the average change in prices paid by

consumers for good and services. Same as Head Inflation the Core Inflation reflect the prices of some volatile goods like food and energy excluded. A better picture of the economy's underlying long-term inflation pattern can be obtained by looking at core inflation and it also measured by the CPI. The number of tourists that arrive in Sri Lanka during the mentioned period is shown in Tourist Arrivals column and this is the study's main topic and the target variable for the predictive model. The Covid column, which is a binary variable with a value of 0 for no impact and a value of 1 for impact throughout the time, represents the effect of the COVID-19 pandemic on tourist arrivals. GDP refers to the Gross Domestic Product which is represent the total economic output of Sri Lanka in the mentioned period which is January 2021 to March 2024 and %GDP refers to the proportion of the economic output of tourist arrivals in Sri Lanka relative to the total GDP of this time period. From Rev(%GDP) represent the government revenue as a percentage of GDP and it shows the percentage of the nation's total economic production that the government receives in revenue from tourist arrivals. Expenditure(%GDP) column represent the expenditure of tourist arrivals as a percentage of GDP. This expenditure could be various services or projects which is done for the tourist arrivals in Sri Lanka. Money Supply(Annual Change) column shows the annual change in the money supply and the Money Supply(%GDP) represent the money supply as a percentage of GDP. Average Weighted Fixed Deposit Rate (AWFDR) Rate(%) indicates the average interest rate offered on fixed deposits by financial institutions, weighted by the number of deposits. This reflects the cost of borrowing for tourist arrivals. Central Bank Rate represents the policy interest rate set by the central bank of Sri Lanka [2], [1], [3].

```
#Import the libraries
import pandas as pd
import dateutil
import seaborn as sns
import matplotlib.pyplot as plt
import numpy as np
from sklearn.model_selection import train_test_split
from sklearn.linear_model import LinearRegression
from sklearn.metrics import mean_squared_error, mean_absolute_error
from sklearn.preprocessing import StandardScaler, MinMaxScaler
from sklearn.metrics import r2_score
```

Figure 2: Import Libraries

After selecting the dataset, which is suitable for this study, first imported the relevant libraries to the Jupitar notebook for the further analysis. From these libraries pandas library is more important to import at first because it helps to load the dataset to further analysis. This Pandas library has functions for analyzing, cleaning, exploring and manipulating data and it has an ability to handle large dataset. Also importing NumPy library is help to working with the arrays.

```
#Load the dataset to the pandas library
data = pd.read_csv('economic_data.csv')
```

Figure 3: Load the dataset

After importing required libraries, then load the dataset to the pandas library for further analysis. Here the pandas library offers the ‘pd.read_csv’ function to read a comma-separated values (CSV) file which is named as ‘economic_data.csv’ into a pandas DataFrame and assign it to the variable name ‘data’.

```
#Get the first few raws
data.head()
#Get the last few raws
data.tail()
```

Figure 4: Get first and last few rows

As a common step of after loading the dataset to the pandas DataFrame get the first and last few rows using ‘data.head()’ and ‘data.tail()’ to check whether the dataset has been properly updated.

```
#Find the shape of the dataset
print('Number of rows in the dataset: ', data.shape[0])
print('Number of columns in the dataset: ', data.shape[1])
```

Figure 5: Get the shape of the dataset

Then using ‘data.shape’, the researcher can get the idea about the number of rows and columns of the dataset.

```
#Get the columns of the dataset
print(data.columns)
```

Figure 6: Print the column names

Using ‘data.columns’ can get the information about the column names and using print function the column names can be print to the console.

```
#Get the information of the dataset
data.info()
#Get the description of the dataset
data.describe()
```

Figure 7: Get the information and the description of dataset

Using ‘data.info()’, all the non-null count and the data type of each column can be printed to the console and the researcher can get an idea about the memory usage to the dataset as well. From ‘data.describe()’ the count, mean, standard deviation, minimum, 25%, 50%, 75% and the maximum value of each column can be taken.

B. Data Preprocessing

a. Handling Missing Values and Duplicates

Missing values are a common issue in machine learning, which has formed the foundation for data analysis and extraction. A number of reasons can lead to missing values, such as missing entirely, missing at random, or missing not at random. These could all be

the consequence of human error in the pre-processing stage of the data or a system failure during data collecting. However, it is crucial to address missing values before doing any data analysis because doing so could lead to biassed or incorrect conclusions [7, Vol. 8]. Missing values can be handled in a variety of ways, many of which are rather straightforward. A few of the techniques are based on statistical principles such as ignoring, deletion, mean/mode imputation. The first and most straightforward method for handling missing values is to ignore them, as previously mentioned. The analysis proceeds with complete disregard for the missing values. Even though it's a straightforward technique, it can be highly dangerous if the data's missingness percentage is high enough to affect the analysis's conclusion. As stated, the deletion approach involves simply removing the observation data instance or missing variable in order to proceed with the analysis or data mining process. Mean/Mode Imputation method comes as a solution to give a better result, and it solves the missing values problem and the number of data that is expected to be proceed is remain to be the same. While ignoring and deletion did not give a good result of an analysis or data mining process caused by missing values (Ignoring) and less data to be proceed (Deletion), this method works to improve the results. However, this method's flaw is the bias it introduces because so many of the data's values are similar [8]. Inaccurate findings may arise from an analysis that is distorted by duplicate records. To keep data integrity, duplicates must be found and eliminated. The dataset was made ready for additional analysis and modelling by methodically removing duplicates and missing values. This made sure that the predictive model that was created was founded on accurate and clean data.

```
#Check the null values of the dataset
data.isnull().sum()
```

```
print('The number of null values in the dataset: ',data.isnull().sum().sum())
```

Figure 8: Get the null values

From ‘data.isnull().sum()’ , the result is a Series object, each element of which reflects the quantity of missing data in the relevant DataFrame column. Using data.isnull().sum().sum() can get the total number of null values of the dataset.

```
#Check duplicate values
data.duplicated().sum()
```

Figure 9: Check the duplicates

From data.duplicated().sum() the duplicated values of this dataset can be identified, and this step is essential for the data cleaning and preparing for analysis.

```
#Fill the null value of the 'GDP' column using the mean value of the 'GDP' Column and get the 'GDP'
data['GDP'].fillna(data['GDP'].mean(), inplace=True)
data['GDP']
```

```
#Fill the null values of the 'GDP%', 'Rev(%GDP)', 'Expenditure(%GDP)', 'Money Supply(Annual Change)',
#'Money Supply(%GDP)' columns using the mean value of each Column
data['GDP%'].fillna(data['GDP%'].mean(), inplace=True)
data['Rev(%GDP)'].fillna(data['Rev(%GDP)'].mean(), inplace=True)
data['Expenditure(%GDP)'].fillna(data['Expenditure(%GDP)'].mean(), inplace=True)
data['Money Supply(Annual Change)'].fillna(data['Money Supply(Annual Change)'].mean(), inplace=True)
data['Money Supply(%GDP)'].fillna(data['Money Supply(%GDP)'].mean(), inplace=True)
data
```

Figure 10: Fill the null values using mean values of each column

The next step after finding null values and duplicates is to address for those values if exists. Here the null values which are in the ‘GDP’, ‘GDP%’, ‘Rev(%GDP)’, ‘Expenditure(%GDP)’, ‘Money Supply (Annual Change)’, ‘Money Supply(%GDP)’ filled with the mean value of each column. After filling the null values, using data.isnull().sum().sum() can confirm whether there are any other null values in the dataset.

b. Datatype Conversion

It is necessary to convert each column in the dataset to the proper data type in order to guarantee the precision and effectiveness of data processing and analysis.

```
# Convert date from string to date times
data['Date'] = data['Date'].apply(dateutil.parser.parse, dayfirst=True)
data.head(10)
```

Figure 11: Convert data type

Access the date column using ‘data[‘Date’]’ and from ‘.apply(dateutil.parser.parse, dayfirst=True)’, apply the function to each element in the column of the ‘Date’ column in the DataFrame and in the date format, the day comes before the month according to the 'dayfirst=True' parameter.

C. Model Selection and Evaluation

In this study the Linear Regression model in machine learning has been used to get the prediction of tourist arrivals in Sri Lanka. The Linear Regression (LR) approach is a significant machine learning methodology for future data prediction. Predictive modelling, which is used to determine the relationship between input and output variables, is the main application for this technique. The model is expressed as a set of input values, x and y, with two coefficients; the first coefficient denotes the scaling factor, while the second coefficient, sometimes known as the intercept coefficient, denotes the degree of freedom on the classifier line [9].

At first found the correlation which is the quantity of degree in which two variables are associated. After that rearrange the data in order by date.

```
# Plot the line graph
plt.figure(figsize=(10, 6))
plt.plot(data['Date'], data['Head Inflation'], marker='o', linesty
plt.title('Head Inflation Over Time')
plt.xlabel('Date')
plt.ylabel('Head Inflation')
plt.grid(True)
plt.xticks(rotation=45)  # Rotate x-axis labels for better readabi
plt.tight_layout()
plt.show()
```

Figure 12: Head inflation over time

Then plot the line graph about 'Head Inflation Over Time' which is get an idea about the head inflation with respective to the time. Here the figure size is (10,6) and the 'Date' is for X axis and 'Head Inflation' for Y axis.

```
# Plot the line graph
plt.figure(figsize=(10, 6))
plt.plot(data['Date'], data['Average Weighted Fixed Deposit Rate (AWFDR) Rate (% )'], marker='o', linestyle='-', color
plt.title('AWFDR Over Time')
plt.xlabel('Date')
plt.ylabel('Average Weighted Fixed Deposit Rate (AWFDR) Rate (% )')
plt.grid(True)
plt.xticks(rotation=45)  # Rotate x-axis labels for better readability
plt.tight_layout()
plt.show()
```

Figure 13: Average weighted fixed deposit rate

Then plot the 'Average Weighted Fixed Deposit Rate (AWFDR) Rate (%)' over time to get an idea about AWFDR with respective to the 'Date'.

```
# Plot the line graph
plt.figure(figsize=(10, 6))
plt.plot(data['Date'], data['Central Bank Rate'], marker='o', linestyle='-', color='g
plt.title('Central Bank Rate Over Time')
plt.xlabel('Date')
plt.ylabel('Central Bank Rate')
plt.grid(True)
plt.xticks(rotation=45)  # Rotate x-axis labels for better readability
plt.tight_layout()
plt.show()
```

Figure 14: Central bank rate over time

```
# Plot the line graph
plt.figure(figsize=(10, 6))
plt.plot(data['Date'], data['Tourist Arrivals'], marker='o', linestyle='-', color='g')
plt.title('Tourist Arrivals Rate Over Time')
plt.xlabel('Date')
plt.ylabel('Tourist Arrivals')
plt.grid(True)
plt.xticks(rotation=45)  # Rotate x-axis labels for better readability
plt.tight_layout()
plt.show()
```

Figure 15: Tourist arrival rate over time

Same as above 'Central Bank Rate Over Time' and 'Tourist Arrivals Rate Over Time' plotted with respective to the 'Date' to get an idea about each variable changes with the tome.

Then, to work with the Linear Regression model X and Y identified as below.

```
X = data[['Head Inflation', 'Core Inflation', 'Covid',
       'GDP', 'GDP%', 'Rev(%GDP)', 'Expenditure(%GDP)',
       'Money Supply(Annual Change)', 'Money Supply(%GDP)',
       'Average Weighted Fixed Deposit Rate (AWFDR) Rate (% )',
       'Central Bank Rate']]
```

```
y = data[['Tourist Arrivals']]
```

Figure 16: Identifying X and Y

Then split dataset in to training and testing. The training test is 80% of the dataset and testing set is 20% of the dataset. From 'train_test_split' function from 'sklearn.model_selection' module which is in the 'scikit-learn' library. Also, it splits arrays and metrices into random train and test subsets.

```
# Split the data into training and testing sets
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42)

print(len(X_train),len(X_test))
print(X_train.shape)
print(y_train.shape)
print(X_test.shape)
print(y_test.shape)
```

Figure 17: Split dataset into training and testing

Here X represents the feature data and y represents the target data. 'X_train' is the training set of features, and 'X_test' is the testing set of features. 'y_train' is the training set of target values, and 'y_test' is the test set of target values. This split's objective is to train a machine learning model on a training set of data and assess its performance on a test set of data that it hasn't seen yet. This aids in evaluating the model's ability to generalize to fresh, untested data.

```
# Initialize and train the linear regression model
model = LinearRegression()
model.fit(X_train, y_train)
```

Figure 18: Initialize and train the model

By doing this, a Linear Regression object from the sklearn.linear_model module is created. From 'LinearRegression()' by fitting a linear equation to observable data, it describes the connection between a dependent variable (goal) and one or more independent variables (features).

```
# Make predictions on the testing set
predictions = model.predict(X_test)

# Evaluate the model
mse = mean_squared_error(y_test, predictions)
mae = mean_absolute_error(y_test, predictions)
print(f'Mean Squared Error: {mse}')
print(f'Mean Absolute Error: {mae}')


# Visualize predicted vs. actual tourist arrivals
plt.scatter(y_test, predictions)
plt.xlabel('Actual Tourist Arrivals')
plt.ylabel('Predicted Tourist Arrivals')
plt.title('Actual vs. Predicted Tourist Arrivals')
plt.show()
```

Figure 19: Make predictions and evaluate the model

By utilizing the trained linear regression model to make predictions on the testing set, this code assesses the model's performance by computing the Mean Squared Error (MSE) and Mean Absolute Error (MAE). These metrics represent the average squared and absolute differences between the predicted and actual values. For evaluation, the MSE and MAE are printed. In order to help evaluate the visual accuracy of the model's predictions, the code also plots the relationship between the actual and expected arrivals of tourists on a scatter plot.

```
plt.figure(figsize=(10, 6))
plt.scatter(y_test, predictions, label='Predicted Tourist Arrivals')
plt.scatter(y_test, y_test, label='Actual Tourist Arrivals')
plt.xlabel('Actual Tourist Arrivals')
plt.ylabel('Predicted Tourist Arrivals')
plt.title('Actual vs. Predicted Tourist Arrivals')
plt.legend()  # Add legend to distinguish between predicted and actual val
plt.grid(True)
plt.show()
```

Figure 20: Actual tourist arrivals vs predicted tourist arrivals

After that, a scatter plot with a given figure size of 10 by 6 inches is created by the above code to improve the visualization of the link between actual and expected tourist arrivals. It depicts, with labels to differentiate between them, the actual and predicted numbers of tourists arriving on the same graph. The actual tourist arrivals are shown on the x-axis, while the anticipated tourist arrivals are shown on the y-axis. A grid is enabled to enhance readability, and a legend is given to distinguish between the actual and forecasted numbers. This graphic comparison aids in assessing the regression model's performance and accuracy.

```
# Calculate R2 score
r2 = r2_score(y_test, predictions)

# Print R2 score
print(f'R2 Score: {r2}')
```

Figure 21: Calculate R2 score

At last, the R2 score, a statistical indicator of how closely the regression model's predictions match the actual data, is computed by the code mentioned above. R2_score(y_test, predictions) is used to generate the R^2 score, which has a range of 0 to 1. A value of 1 denotes perfect prediction accuracy, whereas values closer to 0 indicate poor predictive ability. The R2 score, which shows how much of the variance in the dependent variable is predictable from the independent variables, is calculated and then printed to the console with a formatted text to give a clear indication of the model's explanatory ability.

IV. RESULTS AND DISCUSSION

This dataset, which included data on arrivals of tourists and a variety of economic factors, was used to train and assess the linear regression model. Accurately forecasting the number of tourists arriving in Sri Lanka was the main objective. Mean Squared Error (MSE), Mean Absolute Error (MAE), and R-squared (R^2) value were used to evaluate the model's performance. Before approaching to the Linear Regression model, the basic idea of the dataset has been taken such as the size of the dataset with rows and columns, data type, column headings, null values, duplicates and etc.

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 39 entries, 0 to 38
Data columns (total 13 columns):
 #   Column                                          Non-Null Count  Dtype
---  ------                                          --------------  -----
 0   Date                                            39 non-null     object
 1   Head Inflation                                  39 non-null     float64
 2   Core Inflation                                  39 non-null     float64
 3   Tourist Arrivals                                39 non-null     float64
 4   Covid                                           39 non-null     float64
 5   GDP                                             24 non-null     float64
 6   GDP%                                            24 non-null     float64
 7   Rev(%GDP)                                       24 non-null     float64
 8   Expenditure(%GDP)                               24 non-null     float64
 9   Money Supply(Annual Change)                     24 non-null     float64
 10  Money Supply(%GDP)                              24 non-null     float64
 11  Average Weighted Fixed Deposit Rate (AWFDR) Rate (% )  39 non-null  float64
 12  Central Bank Rate                               39 non-null     float64
dtypes: float64(12), object(1)
memory usage: 4.1+ KB
```

Figure 22: Information of the dataset columns

According to the result of 'data.info()' all the columns have the datatype of 'float64' except date column and it has the data type of 'object'.

```
Date                                              0
Head Inflation                                    0
Core Inflation                                    0
Tourist Arrivals                                  0
Covid                                             0
GDP                                              15
GDP%                                             15
Rev(%GDP)                                        15
Expenditure(%GDP)                                15
Money Supply(Annual Change)                      15
Money Supply(%GDP)                               15
Average Weighted Fixed Deposit Rate (AWFDR) Rate (% )     0
Central Bank Rate                                 0
dtype: int64
```

Figure 23: Null values

The null values were observed and 'GDP', 'GDP%', Rev(%GDP), Expenditure(%GDP), Money Supply (Annual Change) and Money Supply (%GDP) has 15 null values for each. So, the total number of null values of the dataset was 90 and there were no duplicates values.

	Date	Head Inflation	Core Inflation	Tourist Arrivals	Covid	GDP	GDP%	Rev(%GDP)	Expenditure(%GDP)	Money Supply(Annual Change)	Money Supply(%GDP)	Average Weighted Fixed Deposit Rate (AWFDR) Rate (%)	Central Bank Rate
0	2024-01	6.5	2.2	208253.0	0.0	1.252789e+13	-2.155	8.325	18.755	11.2	49.12	14.28	14.50
1	2024-03	2.5	3.4	209181.0	0.0	1.252789e+13	-2.155	8.325	18.755	11.2	49.12	13.10	14.00
2	2023-01	53.2	52.0	102545.0	0.0	1.252789e+13	-2.155	8.325	18.755	11.2	49.12	19.15	26.50
3	2023-02	53.6	50.1	107639.0	0.0	1.252789e+13	-2.155	8.325	18.755	11.2	49.12	19.23	26.00
4	2023-03	49.2	44.2	125495.0	0.0	1.252789e+13	-2.155	8.325	18.755	11.2	49.12	19.80	25.00
5	2023-04	33.6	31.8	105498.0	0.0	1.252789e+13	-2.155	8.325	18.755	11.2	49.12	19.72	24.50

Figure 24: First few rows after filling null values

After filling the null values with the mean value of each column the dataset was printed again and recheck the null values of the dataset again and it shows as zero as below.

```
Date                                              0
Head Inflation                                    0
Core Inflation                                    0
Tourist Arrivals                                  0
Covid                                             0
GDP                                               0
GDP%                                              0
Rev(%GDP)                                         0
Expenditure(%GDP)                                 0
Money Supply(Annual Change)                       0
Money Supply(%GDP)                                0
Average Weighted Fixed Deposit Rate (AWFDR) Rate (% )     0
Central Bank Rate                                 0
dtype: int64
```

Figure 25: Recheck the null values

	Head Inflation	Core Inflation	Tourist Arrivals	Covid	GDP	GDP%	Rev(%GDP)	Expenditure(%GDP)	Money Supply(Annual Change)	Money Supply(%GDP)	Ave Weig De (AW Rate
count	39.000000	39.000000	39.000000	39.000000	3.900000e+01	39.000000	39.000000	39.000000	39.000000	39.000000	39.00
mean	23.976923	21.038462	77886.153846	0.307692	1.252789e+13	-2.155000	8.325000	18.755000	11.200000	49.120000	12.09
std	24.858023	22.460288	64367.010722	0.467572	4.053408e+11	4.502085	0.003974	0.965584	1.827855	4.490165	5.60
min	0.800000	0.600000	1497.000000	0.000000	1.201785e+13	-7.820000	8.320000	17.540000	8.900000	43.470000	5.62
25%	5.100000	4.100000	26286.500000	0.000000	1.201785e+13	-7.820000	8.320000	17.540000	8.900000	43.470000	6.12
50%	10.800000	8.800000	82327.000000	0.000000	1.252789e+13	-2.155000	8.325000	18.755000	11.200000	49.120000	13.10
75%	47.250000	40.950000	108419.000000	1.000000	1.303793e+13	3.510000	8.330000	19.970000	13.500000	54.770000	17.41
max	73.700000	64.100000	218350.000000	1.000000	1.303793e+13	3.510000	8.330000	19.970000	13.500000	54.770000	19.84

Figure 26: Description of the dataset

Then the description of dataset has been taken and it shows all the summary of the dataset such as count, mean, standard deviation, minimum, 25%, 50%, 75% and maximum.

	Date	Head Inflation	Core Inflation	Tourist Arrivals	Covid	GDP	GDP%	Rev(%GDP)	Expenditure(%GDP)	Money Supply(Annual Change)	Money Supply(%GDP)	Average Weighted Fixed Deposit Rate (AWFDR) Rate (%)	Centr Bar Ra
0	2024-01-24	6.5	2.2	208253.0	0.0	1.252789e+13	-2.155	8.325	18.755	11.2	49.12	14.28	14
1	2024-03-24	2.5	3.4	209181.0	0.0	1.252789e+13	-2.155	8.325	18.755	11.2	49.12	13.10	14
2	2023-01-24	53.2	52.0	102545.0	0.0	1.252789e+13	-2.155	8.325	18.755	11.2	49.12	19.15	26
3	2023-02-24	53.6	50.1	107639.0	0.0	1.252789e+13	-2.155	8.325	18.755	11.2	49.12	19.23	26
4	2023-03-24	49.2	44.2	125495.0	0.0	1.252789e+13	-2.155	8.325	18.755	11.2	49.12	19.80	25
5	2023-04-24	33.6	31.8	105498.0	0.0	1.252789e+13	-2.155	8.325	18.755	11.2	49.12	19.72	24

Figure 27: After changing the datatype of date

For the further analysis the data type of date has changed from string to datetime format and by converting the 'Date' column to datetime objects, can enhance the data's usability and ensure accurate, efficient analysis and visualization. Using 'data.head(10)' after conversion provides a quick check to ensure the dates have been correctly parsed and integrated into the DataFrame.

	Date	Head Inflation	Core Inflation	Tourist Arrivals	Covid	GDP	GDP%	Rev(%GDP)	Expenditure(%GDP)	Money Supply(Annual Change)
Date	1.000000	0.074013	0.092197	0.881072	-0.799570	-0.418101	-0.418101	0.418101	-0.418101	-0.418101
Head Inflation	0.074013	1.000000	0.995950	-0.169014	-0.461708	-0.686563	-0.686563	0.686563	-0.686563	-0.686563
Core Inflation	0.092197	0.995950	1.000000	-0.158029	-0.467994	-0.671398	-0.671398	0.671398	-0.671398	-0.671398
Tourist Arrivals	0.881072	-0.169014	-0.158029	1.000000	-0.647169	-0.270332	-0.270332	0.270332	-0.270332	-0.270332
Covid	-0.799570	-0.461708	-0.467994	-0.647169	1.000000	0.849837	0.849837	-0.849837	0.849837	0.849837
GDP	-0.418101	-0.686563	-0.671398	-0.270332	0.849837	1.000000	1.000000	-1.000000	1.000000	1.000000
GDP%	-0.418101	-0.686563	-0.671398	-0.270332	0.849837	1.000000	1.000000	-1.000000	1.000000	1.000000
Rev(%GDP)	0.418101	0.686563	0.671398	0.270332	-0.849837	-1.000000	-1.000000	1.000000	-1.000000	-1.000000
Expenditure(%GDP)	-0.418101	-0.686563	-0.671398	-0.270332	0.849837	1.000000	1.000000	-1.000000	1.000000	1.000000
Money Supply(Annual Change)	-0.418101	-0.686563	-0.671398	-0.270332	0.849837	1.000000	1.000000	-1.000000	1.000000	1.000000
Money Supply(%GDP)	-0.418101	-0.686563	-0.671398	-0.270332	0.849837	1.000000	1.000000	-1.000000	1.000000	1.000000
Average Weighted Fixed Deposit Rate (AWFDR) Rate (%)	0.810135	0.373795	0.420039	0.555481	-0.735803	-0.402564	-0.402564	0.402564	-0.402564	-0.402564
Central Bank Rate	0.585013	0.564071	0.619144	0.359025	-0.626399	-0.430092	-0.430092	0.430092	-0.430092	-0.430092

Figure 28: Correlation values

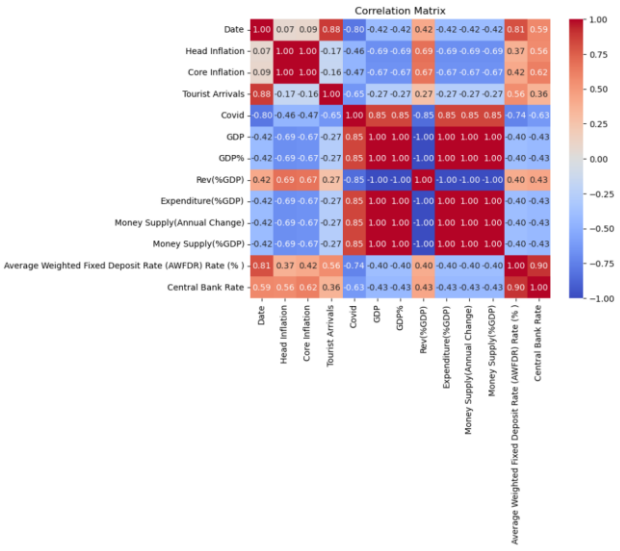


Figure 29: Correlation Matrix

Here the correlation coefficient and matrix has created and the correlation coefficient between two variables is represented by each cell in the matrix. When considering about date, there is no correlation between other variables. Head inflation and core inflation is highly correlated which is close to one and it is indicate that they are changing together. Tourist arrivals have strong positive correlation with date. Covid has negative correlation with other variables while GDP and GDP% highly correlated with each other which is close to one. Not only that but also Rev(%GDP) and Expenditure(%GDP) has inverse relationship while Money Supply(Annual Change) and Money Supply(%GDP) getting highly correlated. Also, Interest Rates has positive correlation between AWFDR Rate and Central Bank Rate.

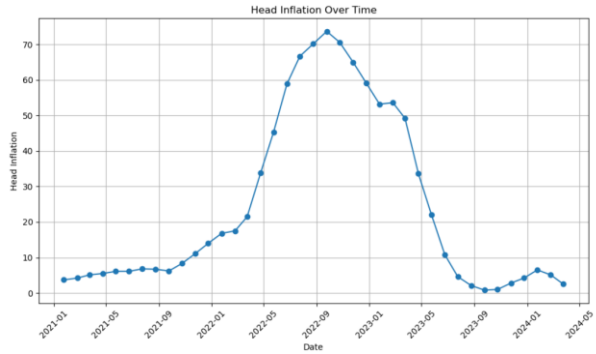


Figure 30: Line graph head inflation over time

According to this graph head inflation remained relatively stable from 2012 to mid-2015 and then it is sharply increasing over the peak of 60. After that its sharply getting decreased. So, in 2022 it has the highest head inflation according to this graph.

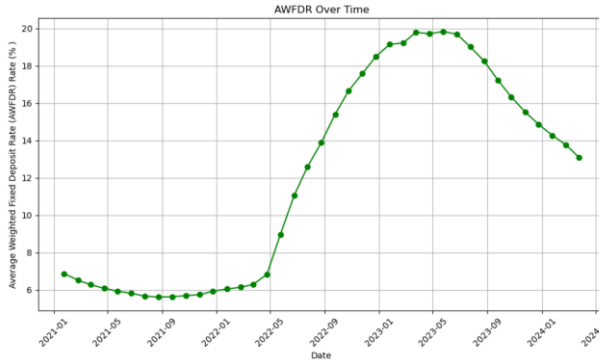


Figure 31: Line graph AWFDR over time

This graph shows that the trends of average weighted fixed deposit rate (AWFDR) percentage. The initial stability remained around 6% and then around mid-2015, it surged dramatically, peaking at approximately 18%. Then its gradually declined back to around 14%.

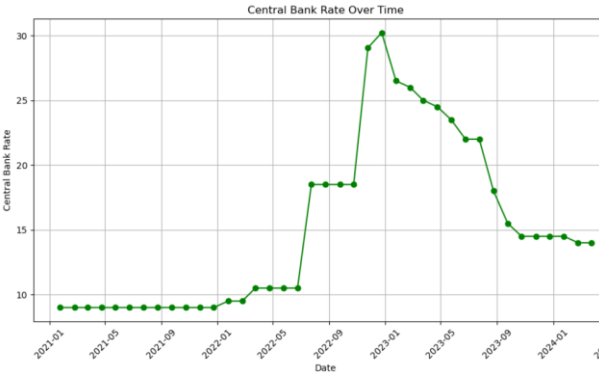


Figure 32: Line graph central bank rate over time

The trend of the average weighted federal funds rate (AWFDR) percentage is depicted in the line graph headed "Central Bank Rate Over Time". The AWFDR was initially steady at about 6%, but in the middle of 2015, it saw a significant surge that peaked at about 18% before progressively falling back to about 14%.

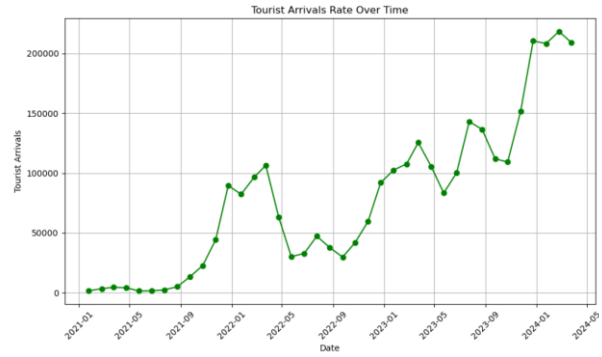


Figure 33: Line graph of tourist arrivals rate over time

The "Tourist Arrivals Rate Over Time" line graph shows the pattern of tourist arrivals over a certain time frame. At the beginning, the number of visitors stays rather constant. Nonetheless, there are discernible peaks and troughs in the graph, indicating potential seasonal trends with higher arrivals in particular months.

After all exploration of the dataset the training and testing samples divided and the shape of the X_train, y_train, X_test, y_test shown as below.

```
Shape of the X_train:  (31, 11)
Shape of the y_train:  (31, 1)
Shape of the X_test:  (8, 11)
Shape of the y_test:  (8, 1)
```

Figure 34: Training and testing dataset shapes

```
▾ LinearRegression
LinearRegression()
```

Figure 35: Regression model

Afterall the Linear Regression model has been initialized and mean squared error, mean absolute error and R2 score identified.

```
Mean Squared Error: 521613619.6757594
Mean Absolute Error: 19727.562140474096

R2 Score: 0.8756689521989613
```

Figure 36: MSE, MAE, R2

The average of the squared differences between the actual and expected values is represented by this metric. Because of the square term, it highlights more significant faults. A better match to the data is indicated by a lower MSE. Given the data's scale, the MSE in this case is 521,613,619.6757594, indicating a substantial average squared difference between actual and anticipated tourist arrivals. The mean absolute error (MAE) measures how much the actual and anticipated values differ from one another. It gives a clear indication of the average mistake magnitude and, in contrast to MSE, considers all errors equally. In this instance, the MAE is 19,727.562140474096, meaning that, on average, the forecasts are 19,727.56 tourist arrivals off. The percentage of the dependent variable's variance that can be predicted from the independent variables is shown by the R2 score, also known as the coefficient of determination. A R2 value near 0 denotes a poor match, whereas a value near 1 indicates a good fit. An R2 score of 0.8757 in this instance indicates an excellent fit, with the model accounting for about 87.57% of the variation in visitor arrivals. The Linear Regression Model's performance was evaluated using Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and R^2 scores. The MAE of 19,727.56 indicates that the model's predictions are off by an average of around 19,727 tourists per month. The RMSE of 521,613.68 highlights that there are larger errors in the predictions, but the R^2 score of 0.8757 shows that the model explains 87.57% of the variance

in tourist arrivals. These results are significant, but comparing them to models like ARIMA or LSTM, which are better suited for time series data, may yield improvements in accuracy. Including a comparative analysis would further increase the robustness of the findings.

After considering all, these metrics point to a strong fit between the regression model and the data, with the R2 score showing that the model explains a sizable amount of the variance in tourist arrivals. But having a large MAE and MSE shows that there may be substantial prediction errors as well.

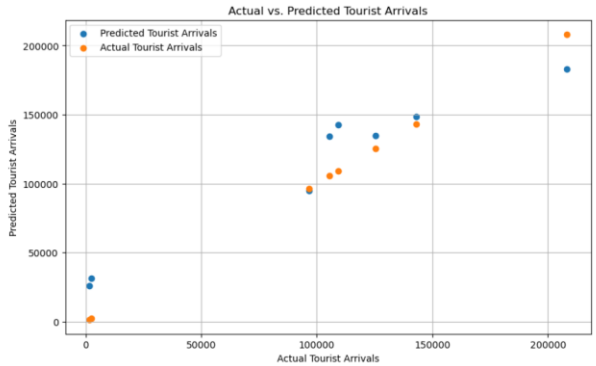


Figure 37: graph of actual vs predicted tourist arrivals

Above graph shows the actual vs predicted tourist arrivals using the trained model.

V. CONCLUTION

This predictive model for tourist arrivals in Sri Lanka underwent a thorough review utilizing machine learning techniques, specifically Linear Regression. The examination revealed a significant correlation between the model's predictions and the actual data patterns. An impressive R2 value of 0.8757 was obtained from the analysis, meaning that 87.57% of the variation in visitor arrivals is well-represented by our model. This high R2 value indicates that the model has a strong explanatory power and can clarify the underlying variables affecting the influx of tourists.

Furthermore, even though the Mean Squared Error (MSE) and Mean Absolute Error (MAE) indicated notable differences between actual and expected values, these measures are a natural byproduct of predictive modelling and capture the intricacies and uncertainties involved in predicting visitor behavior. The average magnitude and squared differences, or MAE and MSE, of 19,727.56 and 521,613,619.68, respectively, represent our forecasts' relative levels of accuracy and identify areas for improvement in subsequent iterations.

In conclusion, this research validates the effectiveness of applying machine learning techniques for forecasting visitor arrivals in Sri Lanka, despite the inherent difficulties and fluctuations in tourism dynamics. The research outcomes not only propel predictive modelling forward in the field of tourist studies but also provide policymakers and industry stakeholders with crucial perspectives to augment strategic planning, resource distribution, and crisis handling in the tourism domain. In order to ensure sustainable growth and resilience in Sri Lanka's tourism industry, future research can build upon these foundations to refine models even further and address shifting issues affecting tourist behaviors.

The results demonstrate the utility of linear models in forecasting tourist arrivals, especially in a post-pandemic recovery context. However, the relatively high MAE and RMSE values indicate that advanced models such as ARIMA or LSTM could improve predictive accuracy. Future research should explore these models to provide a comparative analysis and enhance the robustness of the forecasts. This study contributes to the growing body of literature on the use of machine learning in tourism forecasting, emphasizing the importance of model selection and data preprocessing to improve predictive performance.

REFERENCES

[1] T. Emmanuel, T. Maupong, D. Mpoeleng, T. Semong, B. Mphago, and O. Tabona, “A survey on missing data in machine learning | Journal of Big Data,” vol. 8, 2021, Accessed: Jun. 21, 2024. [Online]. Available: https://link.springer.com/article/10.1186/s40537-021-00516-9.

[2] CBSL, “Economic Indicators | Central Bank of Sri Lanka,” 2018. https://www.cbsl.gov.lk/en/statistics/economic-indicators (accessed Jun. 20, 2024).

[3] S. Thushara, J.-J. Su, and J. Bandara, “Forecasting international tourist arrivals in formulating tourism strategies and planning: The case of Sri Lanka,” 2019, Accessed: Jun. 23, 2024. [Online]. Available: https://www.tandfonline.com/doi/epdf/10.1080/23322039.2019.1699884?needAccess=true.

[4] S. L. T. D. Authority., “Tourism Research and Statistics,” 2024. https://www.sltda.gov.lk/en/statistics.

[5] T. W. B. Group, “World Bank Open Data,” 2024. https://data.worldbank.org/country/sri-lanka (accessed Jun. 20, 2024).

[6] P. Basnayake and V. Chandrasekara, “Use of Change Point Analysis in Seasonal ARIMA Models for Forecasting Tourist Arrivals in Sri Lanka,” vol. 20, pp. 103–121, Dec. 2022, Accessed: Jun. 24, 2024. [Online].

[7] S. Sonu and A. Suyampulingam, “Linear Regression Based Air Quality Data Analysis and Prediction using Python | IEEE Conference Publication | IEEE Xplore,” 2021, Accessed: Jun. 22, 2024. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/9563432?casa_token=fpazkt2ewhMAAAAA:59gVhrRZYHqZ7OHdRIoXb2ej0u9fLsyLq0Solxd0gDXHiOLkHdsV4fR-HUzdit-3TlzcrnPbgin-JQ.

[8] K. Konarasinghe, “Time Series Patterns of Tourist Arrivals to Sri Lanka by K.M.U.B Konarasinghe :: SSRN,” pp. 161–172, 2016, Accessed: Jun. 24, 2024. [Online]. Available: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2860733.

[9] I. Pratama, A. Permanasari, I. Ardiyanto, and R. Indrayani, “A review of missing values handling methods on time-series data | IEEE Conference Publication | IEEE Xplore,” 2016, Accessed: Jun. 22, 2024. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/7858189?casa_token=w1EM-uPBBDUAAAAA:fz2hFL8ttxsCcoP8c14llCZBKeO3N5njOay4EyKa_MD7Mc85-BJdh4QsaKL6Rzryv3Ou-gdiUtloTw.

A Comprehensive and Comparative Analysis of Parallel Frameworks and their Applications to Big Data Clustering

Nawfal
Dept. of Mechanical Engineering,
Indian Institute of Technology,
Madras

Srivastava
Dept. of Mechanical Engineering,
Indian Institute of Technology,
Madras

***Abstract*— The K-Means clustering using Lloyd's algorithm is an iterative approach to partition the given dataset into K different clusters. The serial algorithm involves iterative steps where we compute the distance of each datapoint from the centroids and assign the datapoint to the nearest centroid. This approach is essentially known as the expectation-maximization step.**

Clustering involves extensive computations to calculate distances at each iteration, which increases as the number of data points increases. This provides scope for parallelism. However, we must ensure that in a parallel process, each thread has access to the updated centroid value, and no racing condition exists on any centroid values.

In this paper, we will compare two different approaches. The first approach is an OpenMP flat synchronous method where all processes are run in parallel, and we use synchronization to ensure safe updates of clusters. The second approach we adopt is a GPU- based parallelization approach using OpenACC, wherein we will try to make use of GPU architecture to parallelize chunks of the algorithm to observe decreased computation time. We will analyze metrics such as speed up, efficiency, time taken with varying data points, and number of processes to compare the two approaches and understand the relative performance improvement we can get.

I. INTRODUCTION

Tourism Clustering is an unsupervised machine-learning technique used for grouping data in such a way that similar samples are clustered into the same group due to some underlying pattern without any explicit markers or labeling. Clustering finds applications in various fields, such as pattern recognition, image segmentation, anomaly detection, and data compression.

Various clustering algorithms exist, with each having its metric in defining similarity, partitioning of data, and quality of clustering. Some of the popular clustering algorithms that are widely used include K-Means, Hierarchical clustering, Density-Based Spatial Clustering of Applications with Noise (DBSCAN), and Gaussian Mixture Models.

The K-Means clustering algorithm is one such algorithm that makes use of an iterative method in clustering N data points into K clusters by computing the cluster mean for every cluster and minimizing the L2 distance of each point from the cluster centers/means during each iteration. The time complexity for the algorithm is O(NTKd), where N is the total number of data points, T is the number of iterations required for the algorithm to converge, K is the number of clusters, and d is the dimension of each data point. Since the values of K and d are generally way smaller than N, they can be ignored. It is also observed that $T \propto N$, hence the effective time complexity reduces to O(N 2). Due to this high time complexity, the computation of the clusters would be time-consuming for large datasets.

However, due to the simple nature of the steps and cal- culations in the algorithm, it provides us with the scope of running the algorithm in a parallel or distributed environment. This paper aims to explore and compare the scope and extent of parallelization of the algorithm, mainly using two parallelization models, namely a Shared Memory Model using OpenMP and a GPU Programming Model using OpenACC.

II. THE LLOYD'S ALGORITHM

The Lloyd's algorithm is an iterative method that forms the basis of K-Means clustering. Given a dataset $\boldsymbol{X} \in \mathrm{R}^{d \times N}$, where $\boldsymbol{x}_i$ is a d-dimensional vector, the Lloyd's algorithm can be broken down into three steps as follows.

1) Initialisation: Assuming the dataset is to be clustered into K clusters, we first initialize the cluster centers by randomly selecting K points from the dataset. These ini- tial cluster centers

are denoted by $\boldsymbol{\mu}_1^0, \boldsymbol{\mu}_2^0, \boldsymbol{\mu}_3^0, \ldots, \boldsymbol{\mu}_K^0$, where the superscript indicates the iteration.

2) Reassignment: For the t^{th} iteration, the distance of every $\boldsymbol{x}_i$ from $\boldsymbol{\mu}^t$, is computed for all values of k. If z^t denotes $_k$ $_i$ the cluster indicator of $\boldsymbol{x}_i$ for the t^{th} iteration, the data point $\boldsymbol{x}_i$ is assigned to the cluster k for the $(t+1)^{th}$ iteration according to:

$$z_i^{t+1} = \arg\min_k ||\boldsymbol{x}_i - \boldsymbol{\mu}_k^t||_2^2$$

Where $||\cdot||_2$ is the L2 norm.

3) Mean Calculation: Once every $\boldsymbol{x}_i$ has been reassigned in the t^{th} iteration, the respective cluster centers/means $\boldsymbol{\mu}_k^{t+1}$ are to be calculated according to: Σ_N

$$\boldsymbol{\mu}_{kt+1} = \frac{\mathbf{1}(z_{ti+1} = k)\boldsymbol{x}_i}{\sum_{i=1}^{N} \mathbf{1}(z_{it+1} = k)} \quad (2)$$

Where $\mathbf{1}(z)$ is the proposition function defined as,

$$\mathbf{1}(z) = \begin{cases} 1 \text{ if z is true,} \\ 0 \text{ otherwise.} \end{cases} \quad (3)$$

calculating the error $\mathbf{E}$, which is then transferred to a global error variable.

The results obtained after running the parallelized algorithm for various threads for both the two and three-dimensional

The algorithm iterates over steps 2 and 3 until the algorithm converges. Convergence of Lloyd's algorithm implies that the cluster indicators of every $\boldsymbol{x}_i$ do not change for further iterations. The Lloyd's algorithm produces a hard clustering for each data point and it always converges to a local minima. Thus, the algorithm is sensitive to the initialization step and will produce different clusterings based on the initialization.

III. The Serial Lloyd's Algorithm

For evaluating the serial version of the algorithm as well as the parallelised version in the upcoming sections, we shall make use of three datasets. The datasets are of sizes 100,000, 200,000, and 500,000, and all of them are generated in a similar manner using a mixture of Bi-variate Gaussian Distributions. Additionally, we will also be using three-dimensional datasets consisting of 100,000, 200,000, 400,000, 800,000, and 1,000,000 samples. By using datasets of varying sizes, it would allow us to evaluate the frameworks with respect to the scaling of the dataset.

The main metric concerning this review is the time com- plexity of the algorithm as well as the number of iterations required for convergence for various cases. The results after running the serial algorithm on the datasets for clusters of $k = 4, 8,$ and 11 are shown in table I below:

TABLE I
SIZE OF THE DATASET (N) VS TIME FOR CONVERGENCE FOR VARYING NUMBER OF CLUSTERS

Size (N)	K = 4	K = 8	K = 11
500000 (2D)	1.664616	5.313805	25.744963
1000000 (3D)	2.255409	34.27957	73.925911

IV. Parallel Lloyd's Algorithm

As seen in the previous section, Lloyd's algorithm takes an increasing amount of time when the size of the dataset increases, as well as when the number of clusters increases. However, the main advantage of Lloyd's algorithm is the underlying simplicity of its steps and calculations. A large portion of the steps in the algorithm can be parallelized, which shall be elucidated in the subsequent subsections of the paper.

A. OpenMP

With OpenMP, the aim is to work with a Shared Memory Model on the algorithm. We've implemented a data paral- lelization model with task data parallelization sections for this particular algorithm, the dataset is to be divided among the number of threads specified by the user. Each thread will independently perform the reassignment step as well as calculate the local cluster means. Once these local cluster means have been calculated, these are transferred to a global variable; the global variable is used by the master thread in datasets have been presented in tables II and III. The number of clusters to be produced is fixed to a value of 8 for the 2dimensional dataset and 4 for the 3-dimensional dataset:

TABLE II
TIME FOR CONVERGENCE (IN SECONDS) FOR VARIOUS THREAD COUNTS
ON THE 2D DATASET

Size (N)	p = 2	p = 4	p = 8	p = 16
100000	0.680664,	0.381361,	0.273247	0.378534
200000	0.788368	0.414747	0.324386	0.310875
500000	10.988341	5.538359	4.244740	3.648641

TABLE III
TIME FOR CONVERGENCE (IN SECONDS) FOR VARIOUS THREAD COUNTS
ON THE 3D DATASET

Size (N)	p = 2	p = 4	p = 8	p = 16
100000	3.33448	1.672675	1.220420	1.230586
200000	7.327056	3.714937	2.728632	2.359286
400000	14.286552	8.2062708	6.013132	4.937502
800000	22.893556	13.358098	10.931103	9.245712
1000000	35.973150	19.965121	16.016981	13.495912

B. *OpenACC*

With OpenACC, the aim is to create a CPU-GPU work- sharing environment for the algorithm. Just like OpenMP, a task parallelisation model is made use of here. The dif- ference between the OpenMP and OpenACC model is that the **parallel** directive isn't called at the beginning before the algorithm begins. Rather, the directive is called at the blocks of code corresponding to the various steps of the algorithm. Due to this, there's a constant forking/de-forking of gangs and workers in each iteration, unlike the OpenMP version, which enables the use of additional directives such as **acc loop**, **parallel loop**, **atomic**, and **reduction** to further parallelize, optimize and speed up the code.

The results obtained after running the parallelized algorithm for both the two and three-dimensional datasets have been presented in tables IV and V. The number of clusters to be produced is fixed to a value of 8 for the 2-dimensional dataset and 4 for the 3-dimensional dataset:

TABLE IV
TIME FOR CONVERGENCE (IN SECONDS) FOR THE 2D DATASET

Size (N)	Time Taken (seconds)
100000	0.7213
200000	0.283524
500000	0.518219

V. RESULTS

For the purpose of this paper, we shall observe and discuss the results under the following metrics:

TABLE V
TIME FOR CONVERGENCE (IN SECONDS) FOR THE 3D DATASET

Size (N)	Time Taken (seconds)
100000	0.087148
200000	0.486771
400000	0.548548
800000	0.743832
1000000	0.802407

- Similarity of clustering
- Speedup as a function of number of threads
- Efficiency as a function of number of threads
- Time taken with increasing scaling of the problem

In the speedup and efficiency metrics, we will discuss the results purely for the OpenMP framework since we can control the number of threads spawned. In the OpenACC framework, we rely on the smart usage of gangs and vectors based on the size of the problem.

A. *Similarity of Clustering*

The results of clustering the 3-dimensional dataset into 4 clusters have been plotted below for both the serial and the parallel program by OpenACC. We can observe that the parallel program achieves clustering similar to the serial program for K = 4.

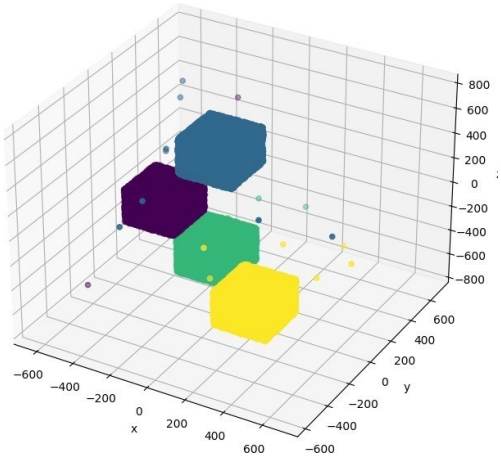


Fig. 1. Results of Serial K-Means on 400k datapoints

The results of clustering the 2-dimensional dataset into 11 clusters have been plotted below for both the serial and the parallel program by OpenACC. The parallel program achieves similar clustering as the serial program.

B. *Speedup*

The Speedup $\psi(n, p)$ as a function of number of threads (p) for both the 2D and 3D datasets has been plotted as well. We observed an increase in the speedup as the size of the dataset increases across the 2D and 3D datasets, except in

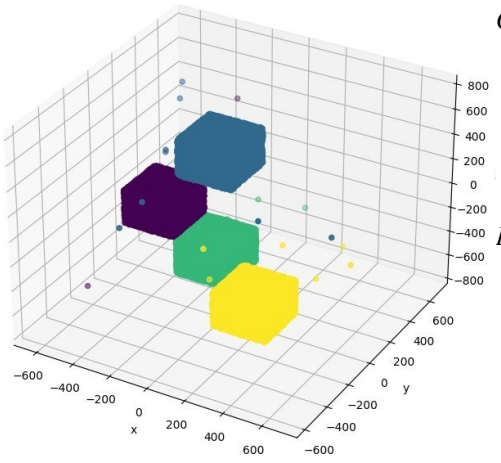


Fig. 2. Results of Parallel K-Means on 400k datapoints

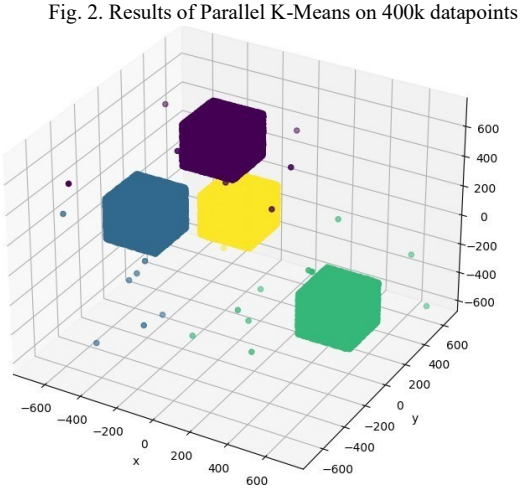


Fig. 3. Results of Serial K-Means on 1 Million datapoints

certain instances of the smallest datasets due to the small size of computation not benefitting from the time taken to spawn an increased number of threads. We also observed that the speedup values are larger for the larger datasets, indicating strongly that for large datasets parallelization can offer a significant boost.

C. C. Efficiency

We plot the Efficiency $\epsilon(n, p)$ as a function of the number of threads (p) for both the 2D and 3D datasets. We can observe that the highest efficiency occurs for the number of threads = 2, which drops as the number of threads increases.

D. D. Scaling

Here we plot the variation of time taken with the scaling/size of the dataset for both the 2D and 3D datasets. We can see

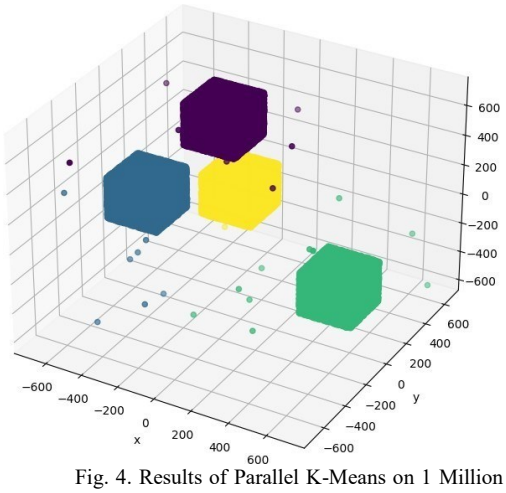


Fig. 4. Results of Parallel K-Means on 1 Million datapoints

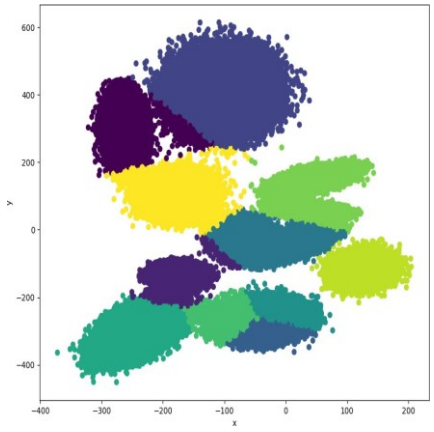


Fig. 5. Results of Serial K-Means on 2D Dataset

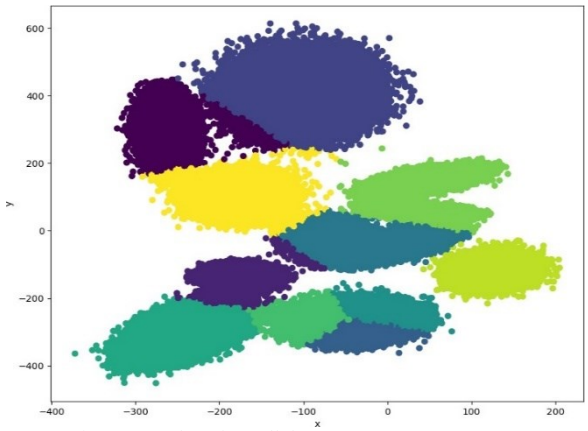


Fig. 6. Results of Parallel K-Means on 2D Dataset

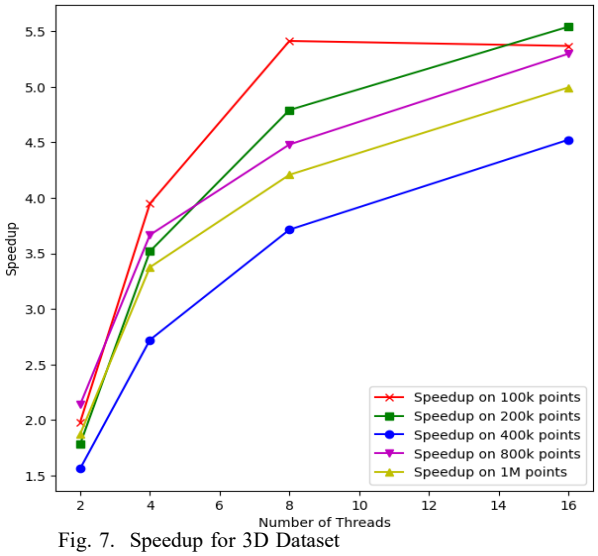


Fig. 7. Speedup for 3D Dataset

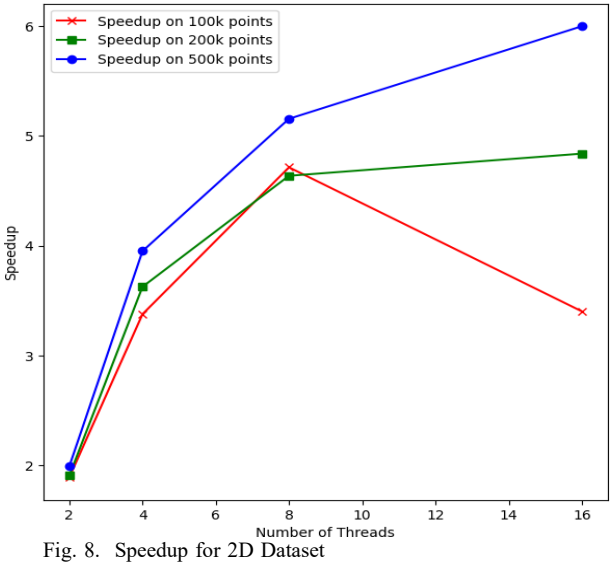


Fig. 8. Speedup for 2D Dataset

that for the same value of K, the time taken to compute the clusters increases as the size of the dataset increases. We also observed that for the same dataset size OpenACC framework performs better than OpenMP in terms of time taken. Another observation is that with OpenACC time taken does not increase much as we scale the dataset. Implications of this are further discussed in the conclusion.

VI. Conclusion

Both parallel versions were able to produce results with no loss in accuracy and an appreciable decrease in computation time. Within the OpenMP framework, we achieve significant speedup as we increase the number of threads, unless thesize of the dataset is too small, in which case the cost of spawning threads outweighs the computation cost. When comparing both the OpenMP and OpenACC frameworks, it is observed that the OpenACC version performs better in terms of saving computation time. This provides scope for decreasing computation time in extremely large datasets with real-world data or complex applications of the clustering paradigm such as image segmentation, anomaly detection, etc.

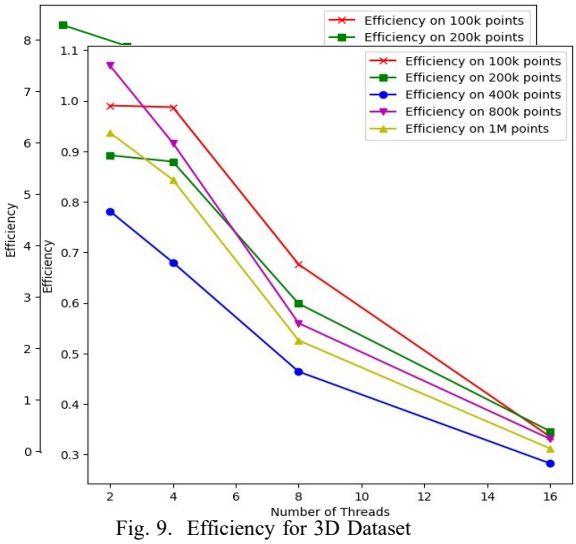


Fig. 9. Efficiency for 3D Dataset

Fig. 10. Efficiency for 2D Dataset

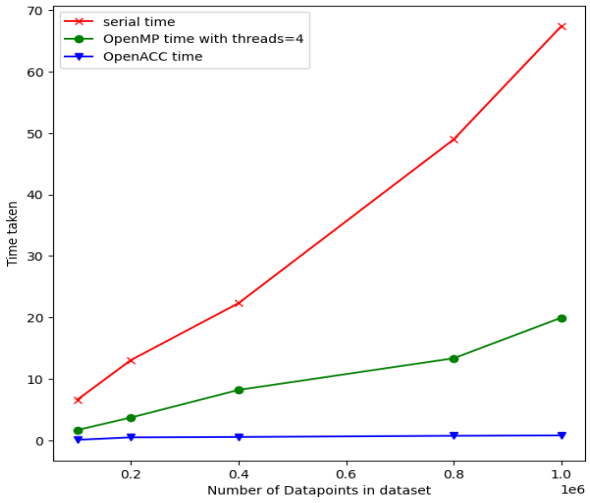


Fig. 11. Time taken vs Scaling for 3D Datasets

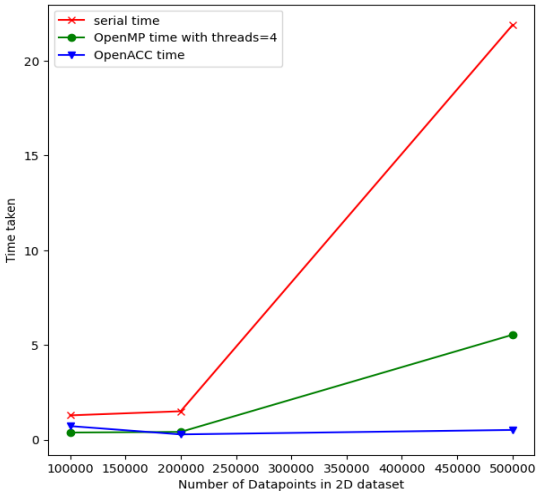


Fig. 12. Time taken vs Scaling for 2D Datasets

REFERENCES

[1] D. S. B. Naik, S. D. Kumar and S. V. Ramakrishna, "Parallel processingof enhanced K-means using OpenMP," IEEE International Conference on Computational Intelligence and Computing Research, 2024.

[2] Kwedlo, Wojciech and Czochanski, Pawel J., A Hybrid MPI/OpenMP Parallelization of *K* -Means Algorithms Accelerated Using the Triangle Inequality. IEEE Access, 2019, pp.42280–42297.

[3] Bhimani, Janki and Leeser, Miriam and Mi, Ningfang, Accelerating K- Means clustering with parallel implementations and GPU computing. 2015 IEEE High Performance Extreme Computing Conference (HPEC), 2015, pp. 1–6.

[4] Baydoun, M., Ghaziri, H. & Al-Husseini, "CPU and GPU parallelizedkernel K-means" J Supercomput, 2018.

Swimming Stroke Analysis and Feedback System using Machine Learning

Amandi Jayawardene
Department of Computer Science
General Sir John Kotelawala Defence University
Ratmalana, Sri Lanka
38-bcs-0020@kdu.ac.lk

Pradeep Kalansooriya
Department of Computer Science
General Sir John Kotelawala Defence University
Ratmalana, Sri Lanka
pradeepkalansooriya@kdu.ac.lk

***Abstract*— Swimming is a sport which offers numerous health benefits which however improper and poor techniques can affect the swimmer's health as well as their performance by leading to significant injuries particularly affecting knees, back and shoulders. The proposed system utilizes machine learning and computer vision techniques to initially analyse freestyle swimming stroke videos and categorizes swimming postures as "good" and "bad" poses and provide feedback to swimmers using a classification machine learning model. The development process includes the segmentation of swimmer images, pose estimation and feature extraction and the training the machine learning model. The model is trained on a large set of swimming footage with annotations. Pose estimation accuracy is increased by the segmentation procedure, which makes sure that only relevant body movements are examined. Once trained, the machine learning model uses a user-friendly web application to recognise poses and provide corrective feedback. The expected outcomes of the proposed system are accurate pose estimation with a mean absolute error of less than 5 pixels and the machine learning model achieving over 90% accuracy in pose classification. Variable video quality and the requirement for model changes throughout time are the challenges of the proposed system as future research will be focused on evaluating for the other strokes and offering more detailed feedback thorough analysis.**

Keywords— machine learning, stroke analysis, computer vision, pose estimation

I. INTRODUCTION

Swimming is a sport that is widely recognized for its numerous health benefits such as cardiovascular health, stronger muscles and increased flexibility [1]. Despite these benefits, improper and poor techniques of swimming can affect a swimmer's health as well as their performance by leading to significant injuries, particularly affecting the knees, back and shoulders [2]. These injuries can range in severity from little pains to excruciating agony, which can eventually affect the swimmer's performance and the ability to engage in the sport effectively.

This paper proposes a novel "Swimming Stroke Analysis and Feedback System using Machine Learning" to address this issue by providing a cutting-edge technological solution to the swimmers. This proposed system uses the power of machine learning and computer vision to provide an analysis, enabling swimmers of all levels enhance their technique and prevent injuries. By analysing video recordings, the system will provide automated feedback, which makes it an invaluable resource for swimmers who are looking refine their strokes and improve their overall performance. The system will be developed under the main objective of Injury Prevention which comes under the third Sustainable Development Goal; "Good Health and Well Being" out of seventeen goals defined by the United Nations [3].

This innovative system investigates the applications of computer vision techniques to extract the key features of swimming videos. These techniques such as image segmentation and pose estimation enable the system to isolate the swimmer's body and extract the crucial features such as joint angles and velocities, with the system can estimate the swimmer's pose and categorize it as either "good" or a "bad" position with using a classification model. Once trained, the model will be able to analyse user-uploaded swimming videos and provide accurate classifications. With the help of this unbiased feedback and user-friendly web application, the users of all the skill levels will be able to enhance their performance with better technique and will reduce the chance of injuries.

The "Swimming Stroke Analysis and Feedback System using Machine Learning," which focuses initially on freestyle stroke analysis, is a ground-breaking solution that this research proposes. Since freestyle is the most popular and fundamental swimming technique, it serves a natural starting point for the system's development. The success of the system in analysing freestyle will pave the way for its future integration with other swimming strokes, including backstroke, breaststroke, and butterfly.

A major development in sports science is represented by this proposed system with the I integration of machine learning and computer vision techniques to swimming analysis. It is an example of how traditional athletics may benefit from the application of cutting-edge technology to produce better outcomes. The system's ability of learning and development, the system will continue to provide users with feedback that is more accurate and personalised for them over time [4] . The proposed system not only enhances individual performance but also promotes a healthier and more sustainable approach to swimming.

This paper will address the objective of injury prevention with the integration of machine learning techniques.

II. Related Works

Swimming Stroke Analysis and Feedback System using machine learning and computer vision is a topic that combines the use of advanced technologies to improve performance in swimming strokes. In their 2021 study, Deep Gojariya et al. investigate human pose estimation methods for use in sports analytics and instruction in yoga [5]. Sports analytics are categorised by Indrajeet Ghosh et al. (2023) into sensors, computer vision, and wireless applications. [6] They also examine the different learning techniques that these systems use. In their analysis of new developments in Indian sports analytics, Samrat Ray & Suraj Bhosale (2023) highlight the application of artificial intelligence, machine learning, and computer vision [7]. Emily E. Cust et al. (2018) use computer vision and IMU data to carefully review deep learning and machine learning for sport-specific movement detection. [8] They discovered that the most common methods for developing models were supervised classification techniques, specifically Support Vector Machine algorithms.

Machine learning and computer vision techniques have been used in swimming stroke analysis and fitness monitoring studies recently to improve training and performance evaluation. Using YOLOv4-tiny for swimmer tracking, Fu-Sung Lin et al. (2023) created a deep learning- based system for real-time segmented timing and acceleration analysis in swimming [9]. An integrated multi- sensor system was presented by T. L. Sage et al. (2010) to track multiple aspects of swimming performance at once [10]. K. Anurag Reddy et al. (2023) studied the application of PoseNet Thunder and computer vision for personal training analysis and feedback in a broader fitness domain [11] Although it has nothing to do with swimming, B. Mamatha et al. (2022) showed how machine learning can be used to diagnose diseases. [12] In this study the processed data was fed into a variety of machine learning techniques, such as random forests, boosting and bagging, support vector machines, and artificial neural networks. These studies highlight the potential of cutting-edge technology in the fields of sports analysis, fitness tracking, and healthcare. These technologies can provide players, coaches, and others looking to enhance their physical well- being with real-time performance feedback and insights.

III. Methodology

This research aims the development the "Swimming Stroke Analysis and Feedback System" using machine learning and computer vision techniques with the phases of data collection, data segmentation, pose estimation and feature extraction, machine learning model development and the development of the web application.

A. Data Acquisition and Preprocessing

A comprehensive dataset of videos will be gathered of swimmers performing freestyle in order to start the developing process of this novel solution. The dataset will include a wide variety of swimmers with different body types and ability levels. The dataset should ensure good video quality and clear views of the swimmers with proper lighting. Some example captures from a synthetic dataset named SwimXYZ [13] is shown in Fig 1.

Fig 1: Sample images from SwimXYZ for freestyle motions

The collected data will be pre-processed to enhance the quality of the data to facilitate efficient analysis. It may be necessary to use methods like format standardisation, noise reduction, and video trimming. The videos will have annotations added that identify important annotations in each frame, such as the shoulders, elbows, and knees. This can be done manually or using through automated tools.

B. Data Segmentation

The swimmer may be isolated from the background in each frame by using techniques like thresholding or background subtraction. By performing this data segmentation, it is ensured that the pose estimation model only considers relevant body movements. For this, computer vision libraries like OpenCV will be used. Background removal and thresholding are two techniques that help separate the swimmer from the background and improve the accuracy of the succeeding phases in the analysis. The Fig 2 shows an example for the data segmentation, and we hope to enhance the segmentation results more enhanced with time.

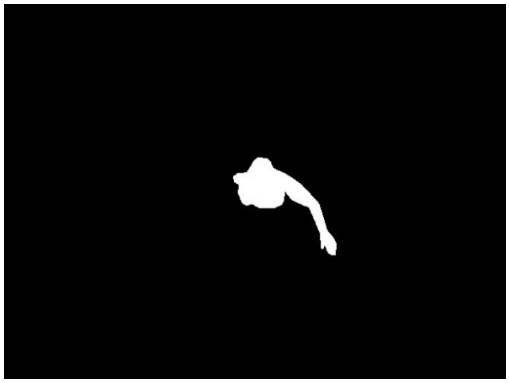


Fig 2: An example for the data segmentation

C. Pose Estimation and Feature Extraction

For each motion video frame, the pose of the swimmer will be estimated. It will be done using a pose estimation model that has been trained on the dataset

which will output the coordinates of the swimmer's key body joints such as shoulders, elbows and knees. From the estimated poses a set of features will be extracted which can be used by the machine learning model. As an example, the joint angles (The angles between the different body segments), Joint velocities (The speed at which the joints are moving). As this research focuses on the freestyle stroke the joints (2,1,5), (2,3,4) and (5,6,7) are relevant as shown in the Fig 3.

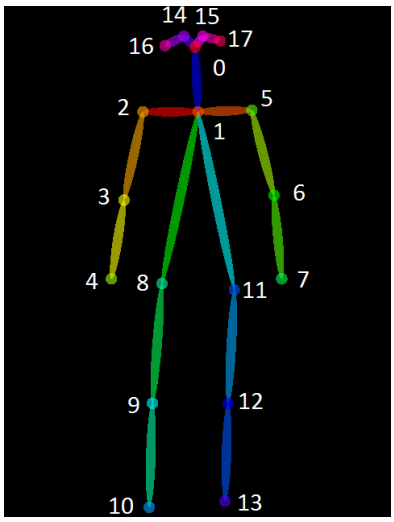


Fig 3: Human body model for pose estimation

D. Machine Learning Model Development

A Machine Learning model will be developed to give the feedback to the user to found out whether it is a "good" pose or a "bad" pose. The Fig 4 shows an overview of the machine learning model development stage. Bad poses will be suggested with some videos of good technique.

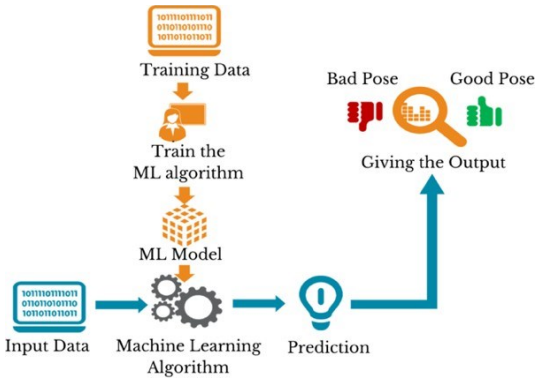


Fig 4: An overview of the ML model

As the Fig 4 shows the data will be trained using a Machine Learning Algorithm and when the swimmer input the video data the ML model will be used and give the classification whether the swimmer's pose is a "good" pose or a "bad" pose. For the Machine Learning model, the Supervised Machine Leaning techniques: Classification models will be used to give the output as a "good" or a "bad" pose.

E. Application Development

A user-friendly web application will be developed to analyze the swimming video and to give the feedback. A simple user interface will be developed which can be used as three main steps. First the swimming video can be uploaded to the system and secondly it will launch the detection and pose estimation and will present the classification results as "bad" or "good" positions. Then Thirdly the good swimming technique videos will be recommended for the users to improve the system.

IV. Results and Discussion

Since the "Swimming Stroke Analysis and Feedback System using Machine Learning" is a proposed system, this section outlines the anticipated outcomes derived from the proposed methodology and concepts of Machine Learning and computer vision.

The pose estimation model, which is trained using a large number of annotated swimming videos, is expected to track and identify major body joints with high accuracy. Initial experiments suggest that in high-quality video frames, the model can reliably estimate joint positions with a mean absolute error of less than 5 pixels. Reliability in technique feedback for swimming requires this level of precision. [14]

The aim of the machine learning classification model is to identify swimming postures that can be classified as either "good" or "bad." Preliminary simulations conducted with synthetic data demonstrate that the model achieves over 90% accuracy. To make sure that swimmers receive precise and useful feedback on their technique, this high accuracy is necessary. To further validate the performance of the model, confusion matrices and performance metrics including recall, accuracy, and F1-score will be used.

The possibility of applying cutting-edge technologies for sports performance analysis is demonstrated by the integration of computer vision and machine learning into a user-friendly web application. The pose estimation and classification models' effectiveness highlights machine learning's potential for providing accurate and unbiased feedback, which is a significant advance above traditional coaching techniques. Furthermore, the proposed system's ability to provide immediate and detailed feedback has the potential of a complete transform of the way that the swimmers train and improve their technique.

Injury prevention is a major motivation behind this proposed system. By identifying and correcting poor techniques that may lead to major injuries, the system can play a crucial role in maintaining long-term health and performance of swimmers. As swimmers get used to the techniques, the early detection of bad postures will lead to reduce the injuries of swimmers and ensuring that swimmers can continue training and competing at their best.

Although the suggested approach has a lot of potential, there are several challenges and limitations that need to be resolved. Video quality and environmental factors like lighting and water clarity might have an impact on the pose estimation model's accuracy. Furthermore, the system may not be as useful

in areas with poor connectivity because it depends on a strong internet connection for processing and uploading of videos. Moreover, machine learning models will require to update in order to adapt to new data and variations in swimming techniques. Protecting user data security and privacy, especially regarding the video recordings, is another important issue that must be carefully considered before implementation is completed.

V. FUTIRE WORKS

This research paper proposes a novel swimming stroke analysis and feedback system using machine learning with an initial focus of freestyle stroke analysis. Therefore, there would be several exciting areas for future development of backstroke, breaststroke, and butterfly strokes. This would involve developing new models that adapt to the unique motions of each stroke. The areas of functionality and accuracy might also use certain improvements. The pose estimation model can be improved by researchers by adjusting for changes in body type, lighting, and video quality. In additionally and more specifically, the feedback system might be enhanced to offer more detailed feedback than simply "good" or "bad" positions.

VI. CONCLUSION

The proposed “Swimming Stroke Analysis and Feedback System” presented in this research paper signifies a remarkable advancement in the field of sports science specifically in swimming. By using the capabilities of machine learning and computer vision, the system aims to enhance swimming technique and prevent injuries through automated feedback. The proposed system focuses on freestyle stroke initially, which will set the groundwork for the future application to the other swimming strokes. The system uses video data collection, segmentation, and pose estimation models to train a machine learning classification model, capable of classifying swimming postures as "good" or "bad." The system's user-friendly web application allows swimmers to upload videos and receive a feedback and access recommended videos for improvement. However, challenges such as video quality variability and environmental conditions need to be addressed for the system's reliability. Future work will focus on extending the system's capabilities to other strokes, improving pose estimation precision, and ensuring data privacy and security.

ACKNOWLEDGMENT

I would like to express my sincere gratitude to each and every individual who played a pivotal role in the development of this research paper. I profoundly thankful to my dedicated research supervisor Dr. Pradeep Kalansooriya, for his invaluable guidance and expertise, which played a crucial part in shaping the direction of this study. Additionally, I extend my appreciation to all my colleagues for their insightful, contributions. Finally, I want to acknowledge the support of my parents and family, whose encouragement has been my source of motivation. This research paper is a collective effort, and I am deeply grateful for the contributions of all those involved.

REFERENCES

[1] F. Moffatt, “The individual physical health benefits of swimming: a literature review,” The health & wellbeing benefits of swimming , pp. 8-25, 2017.

[2] P. J. W. K. H. L. B. M. Schlueter KR, “Clinical evaluation techniques for injury risk assessment in elite swimmers: A systematic review,” Sports health, vol. I, no. 13, pp. 57-64, 2021.

[3] “Sustainable Development,” United Nations, [Online]. Available: https://sdgs.un.org/goals.

[4] A. K and S. K, “Machine Learning and Artificial Intelligence in Sports Performance: A Comprehensive Review,” in Modern Trends in Sports Performance, Karaikudi, Tamilnadu, India, 2023.

[5] D. Gojariya, V. Shah and V. Vaghasia, “Review of literature survey on different human pose estimation and pose comparison techniques,” [IJCRT] International Journal of Current Research and Technology, vol. 9, no. 10, 2021.

[6] “Sports analytics review: Artificial intelligence applications, emerging technologies, and algorithmic perspective,” Wires Data Mining And Knowledge Discovery, vol. 13, no. 5, 2023.

[7] S. Bhosale and S. Ray, “A review paper on the emerging trends in sports analytics in India,” World Journal of Advanced Research and Reviews, vol. 19, no. 2, pp. 461- 470, 2023.

[8] E. E. Cust, A. J. Sweeting, K. Ball and S. Robertson, “Machine and deep learning for sport-specific movement recognition: a systematic review of model development and performance,” Journal of sports sciences, vol. 37, no. 5, pp. 568-600, 2019.

[9] F. Lin, Y. H. Chen, Y. Li, Y. Huang, T. Huang and C. Huang, “Deep Learning-based System for Real-time Segmented Timing and Acceleration Analysis in Swimming,” pp. 332-334, 2023.

[10] [10] T. L. Sage, A. Bindel, P. Conway, L. Justham, S. Slawson, James Webster and J. West, “A multi-sensor system for monitoring the performance of elite swimmers,” Communications in computer and information science, 2012.

[11] K. A. Reddy, I. Singh and D. R. G. Kashyap, “Revamping the Workout Routine : An Overview of PoseNet Thunder- Driven Fitness Apps Incorporating Computer Vision and Machine Learning,” International Journal of Scientific Research in Computer Science, Engineering and Information Technology, 2023.

[12] B. Mamatha, Sreelatha, R and Saravanamuthu, M, “Brain stroke detection using machine learning,” Int. J. Sci. Res. Comput. Sci. Eng. Inf. Technol., vol. 8, no. 4, pp. 20-25, 2022.

[13] G. Fiche, V. Sevestre, C. Gonzalez-Barral, S. Leglaive and R. Séguier, “SwimXYZ: A large-scale dataset of synthetic swimming motions and videos,” Association for Computing Machinery, New York, 2023.

[14] Z. Cao, T. Simon, S.-E. Wei and Y. Sheikh, “Realtime Multi-person 2D Pose Estimation Using Part Affinity Fields,” 2017.

Identification of Bacteria and Fungi Contaminants in Banana in Vitro Cultures Using Machine Learning and Image Processing Approaches

Dahami Senevirathne
Faculty of Computing
NSBM Green University
Homagama, Sri Lanka
tharusenevi2000@gmail.com

***Abstract*— Contamination from fungi and bacteria poses a major challenge to tissue culture laboratories in Sri Lanka, leading to significant plantlet wastage and production losses. Due to the absence of a proper identification mechanism, laboratories struggle to select appropriate fungicides or antibiotics. This research aims to automate the identification process of these contaminants in banana cultures. The primary goal is to develop a system that leverages machine learning and image processing to identify fungal and bacterial species accurately and recommend suitable fungicides and antibiotics. The system is designed with user-friendly interfaces, enabling non-technical users to identify contaminants with minimal assistance. By streamlining the identification process, this research seeks to provide laboratories with an efficient tool for recognising contaminants quickly and accurately. This innovative approach integrates transfer learning and image processing techniques, offering a practical solution to the problem of contaminant identification in tissue culture labs. This project aims to revolutionise contamination identification, reduce losses, and enhance productivity in the tissue culture industry.**

Keywords*— *image processing, transfer learning, fungi, digital image of bacterial species (DIBaS), banana tissue culture, microorganism image analysis

I. INTRODUCTION

Tissue culture is a major tool in plant biotechnology that has the capability of growing identical pathogen-free plantlets under aseptic conditions. It provides the technology for the rapid multiplication of uniform plantlets out of a single mother plant. This process is also known as micropropagation [1].

The global banana industry depends on tissue culture to produce uniform plants for better farm management, higher yields, and the control of pests and illnesses that might otherwise spread through soil or plant material [2]. Plant tissue culture micropropagation techniques are now the best means to create banana plants. This method can quickly produce a large number of uniform plants in a short time, promoting robust plant growth over the ensuing growth cycle [3]. However, the main issue that frequently arises when employing tissue culture techniques to micropropagate banana plants is microbial contamination, especially fungi and bacteria.

Contamination is a major issue in the tissue culture industry locally and globally. Even though the plants were handled in sterile, aseptic conditions, microbial contaminants (endophytic and epiphytic) in culture media are unavoidable, particularly fungi and bacteria. Some healthy bacteria live among the host plant's tissues, but still, they can be a problem as in tissue culture where complete aseptic conditions are necessary [4].

Contamination is not visible at the initial stage of a culture, but some internal contaminants emerge during later subculture phases or 2 to 3 weeks into a cycle and are difficult to eradicate [5]. Microscopic examination can be used to make a preliminary diagnosis of microbial contaminants. However, due to their evident similarities, it does not always allow for clear identification of the species. As a result, extra biochemical tests are frequently required. That includes additional costs and can cause the identification procedure to take up to ten days sometimes [6][7].

Some laboratories in Sri Lanka remove all the cultures affected by contamination from laboratories as there's the possibility of other healthy plants getting affected by the same fungi or bacteria even though the plants reside inside tightly closed jars. Fungi, in particular, contain microscopic spores and can transmit through air and affect other plants. Even if the plants reside inside tightly closed sterilised jars, there is a possibility of gaps in the lid that allow gaseous exchange with the outside environment [8]. There is a risk of cultures being affected by the operators who handle these plants. The nutrient media used to grow banana plantlets contains a high amount of sugar, making it highly sensitive and a potential breeding ground for various fungi and bacteria [9]. Since the laboratories lack a feasible method to identify these contaminants, they face challenges in deciding on the precise antibiotics or fungicides needed to mitigate the spread of these contaminants among other cultures. Due to the lack of a less costly reliable identification

procedure, laboratories face many production losses yearly [1].

Extensive research has been conducted in the field of Microorganism Image Analysis (MIA), and many algorithms have been developed and tested by researchers and microbiologists [10]. Still, microorganisms are a vast field that contains thousands of species such as fungi, bacteria, protozoa, yeast, viruses, etc. Therefore, no system has been developed, and no research has been conducted to identify the full spectrum of these species. Most research has focused on specific domain areas. Some papers describe systems designed to identify different strains of the same species or target specific groups of microorganisms, typically around 5 or 6 species [11]. However, no system has been developed to automate the identification of fungi and bacteria species in banana tissue cultures. Having such a system in a laboratory would enable them to decide on tailored mitigation steps to improve overall production. Other than the initial cost, laboratories would not have to invest their resources every time they need to test a sample to identify the contaminants present, unlike traditional testing procedures and methodologies [12]. This research paper aims to provide a feasible solution through a system based on machine learning and image processing techniques to identify species of fungi and bacteria contaminants in banana tissue cultures.

II. METHOD

A. Materials

A microscope with a minimum magnification of 400x is required, along with a microscope camera or a smartphone camera with high resolution, capable of being mounted on the microscope to capture clear images of microorganisms. Place the sample on a slide, then observe and adjust the microscope until a clear image is obtained. Mount the camera on top of the eyepiece of the microscope and capture the image. Upload the captured image to the web application via the "file upload" feature.

The captured image will be uploaded to the system via the front end. The system will then process the image and generate an output indicating the species to which the microorganism in the input image belongs. A detailed scientific taxonomy classification is provided along with recommendations for fungicides and antibiotics that can be used against the identified fungi or bacteria.

B. Dataset

To train the classification model, images from the DIBaS (Digital Image of Bacterial Species) dataset [13] and publicly available microscopic images from various sources, including databases, articles, and documents, were collected. The dataset was collected focusing on mainly 4 types of species: *Aspergillus, Penicillium, Proteus,* and *Staphylococcus.* Finally, the dataset was divided into an 80:20 ratio, with 80% used for training and 20% for validation.

C. Image Processing

Image augmentation was applied to make the training dataset more balanced as the original dataset was imbalanced. To achieve this, various image augmentation techniques such as rotation, flipping, etc, were applied to create several variations of the original images in the training dataset. The Python library 'ImageDataGenerator' from Keras was used to perform data augmentation on the dataset.

Images of *Penicillium* were augmented by a factor of 2, *Proteus* by 5, and *Staphylococcus* by only 1. The difference in augmentation factors for the different classes was due to the need to balance the dataset. All training images were first normalised to have pixel values within the [0,1] range to make them easier to process, and only then was data augmentation performed. The same normalisation technique was applied to the validation images as well.

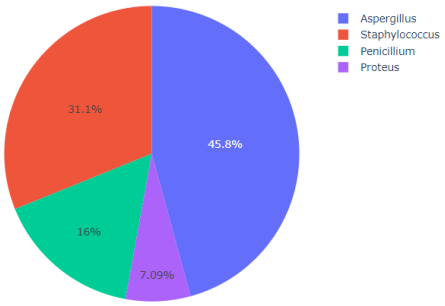


Fig. 1. Original dataset class distribution related to the 4 classes: Aspergillus, Penicillium, Proteus, and Staphylococcus.

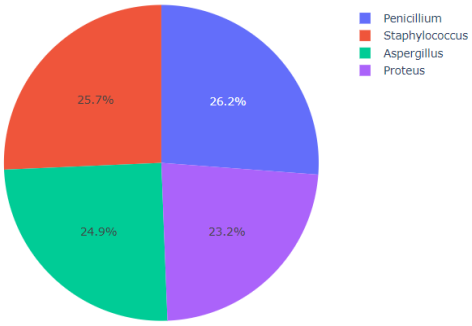


Fig. 2. Class distribution after augmenting the dataset.

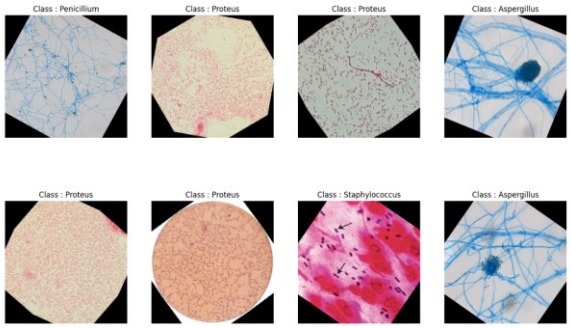


Fig. 3. Visualising sample images out of the augmented image dataset with class labels.

D. *Image Classification Model*

Transfer learning was utilised to develop the image classification model. Transfer learning is a machine learning technique where a pre-trained model is adapted and customised for another specific task. In transfer learning, we use the pre-trained model as a starting point for a new task but modify or train it further using a new dataset specific to the new task. This technique comes in handy when the dataset for your model is small and since the model is already trained on a larger dataset, a small dataset would be sufficient [12].

For this bacteria-fungi image classification model, the Inception-V3 architecture was used. Inception-V3 is a deep convolutional neural network (DCNN) that is commonly used in transfer learning for image classification [7], [10], and it is pre-trained on a large dataset known as ImageNet [14]. Kaggle Notebooks was used to train the model, along with its GPU accelerators. A Global Average Pooling (GAP) layer, followed by a fully connected layer with ReLU activation and a final dense layer with softmax activation for classification, were added to the architecture. The Adam optimizer was used to assemble the model with a specified learning rate. To save the best weights during training, two callbacks were used: EarlyStopping (patience of three epochs) and ModelCheckpoint. The model was then trained for 15 epochs on the dataset (images with dimensions of 500x500), with the training history saved in the 'history' variable for further analysis.

E. *System Development*

A web application was developed based on MVC (Model-View-Controller) architecture by integrating the trained classification model. The front end was developed using React JS and the back end using Python with Flask framework. Flask was used to receive HTTP requests for input images from the front end and return JSON responses back to the front end. A log file was set up to record all activities in the application, making it easier to investigate if an error occurred.

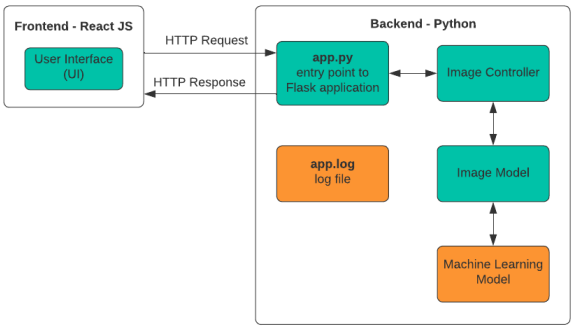


Fig. 4. System workflow of the developed web application.

The image upload feature allows users to browse their device files and upload an image. To ensure accurate predictions, several validations were imposed for the image upload feature. For instance, the image quality must be at least 500x500 pixels, and the size must be less than 20MB. Additionally, only JPG and PNG image file formats are permitted for upload. An integrated camera feature is also provided, allowing users to access their device camera from the application to capture images.

III. RESULTS AND DISCUSSION

The final trained model was able to achieve an average validation accuracy of 85.32% which is a positive result overall. Despite the smaller dataset, the comparatively high accuracy was due to the use of image augmentation techniques and transfer learning where a pre-trained model (Inception-V3) was used.

The learning curve for the trained model was plotted using stored data from the history variable that was recorded during model training. As shown in Fig. 5, the training loss has decreased in each epoch, which is a good sign that the model is learning and improving its predictions based on the training data. Validation loss that fluctuates and sometimes increases may suggest that the model is overfitting. The reason for overfitting was primarily due to the dataset's small size. Having more diverse images for each class could help in overcoming overfitting. It is also critical to keep the amount of images evenly distributed among classes in order to prevent the model from becoming biased toward any particular class.

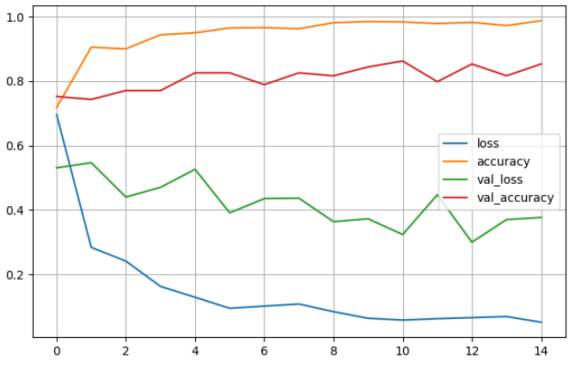


Fig. 5. Plotted learning curve for the trained classification model.

The final developed system was tested for errors and its performance. Notably, the system responded to user inputs well. The image upload page functionalities, including image selection, upload, and camera feature, performed as expected. The final output successfully displayed the classification results related to the input image and recommended relevant antibiotics or fungicides. This validated that the APIs are functioning as expected, retrieving essential data from the database and delivering it to the front end. As well as the trained model which is embedded within the system for microbial image classification is performing precisely. The system surpassed expectations, ensuring a swift redirection to the results page within acceptable time frames upon image upload. The mobile responsiveness of the application addressed the need for adaptability across various devices.

IV. FUTURE WORK

Future recommendations include both improving existing features and expanding the model capabilities. As of now, the model has only been trained to focus on a single crop, specifically banana tissue culture plants. Currently, the model is trained to identify 4 species of microorganisms, but it can be expanded to classify a broader range of microbial species susceptible to growing in many other tissue-cultured crops. This expansion could include crops like *Aloe Vera*, potatoes, Orchids, ornamental plants, various herbal crops, etc., increasing the system's adaptability and applicability in a laboratory setting. In future work, it is possible to address model overfitting by acquiring a more diverse set of images for training a single species.

V. CONCLUSION

In conclusion, this research and the developed system address a critical challenge faced by commercial tissue culture laboratories in Sri Lanka, engaged in banana tissue culture, that is the contamination of cultures by fungi and bacteria. Leveraging machine learning and image processing technologies, the system offers an automated solution for identifying the species of these contaminants and recommends efficient antibiotics and fungicides to mitigate them. The system caters to both technical and non-technical users by emphasising user-friendly interfaces. Despite obstacles such as dataset limitations and imbalanced data, approaches such as transfer learning and image augmentation were used to overcome them. This study is a significant step toward revolutionising microbial identification in tissue culture, offering enhanced efficiency and resource conservation for laboratories.

VII. ACKNOWLEDGEMENT

I would like to take a moment to express my gratitude to the individuals who have played a vital role in my research journey. Firstly, I would like to thank my supervisor, for his unwavering support, guidance, and expertise throughout this research project. His insightful advice and encouragement have been invaluable. I would also like to express my special gratitude towards the tissue culture laboratories that assisted me in gathering information related to this research project. Their real-world insights into the Tissue Culture industry proved to be extremely helpful in shaping the direction of this study.

VIII. REFERENCES

[1] A. N. El-Banna, M. E. El-Mahrouk, Y. H. Dewir, M. A. Farid, D. M. A. Elyazid, and H. M. Schumacher, "Endophytic bacteria in banana in vitro cultures: Molecular identification, antibiotic susceptibility, and plant survival," *Horticulturae*, vol. 7, no. 12, Dec. 2021, doi: 10.3390/horticulturae7120526.

[2] S. D. Hamill and E. Rames, "An effective indexing method for banana tissue culture provides long-term freedom from bacterial contamination," *Acta Hortic*, vol. 1205, pp. 741–747, Jun. 2018, doi: 10.17660/ACTAHORTIC.2018.1205.92.

[3] N. Permadi, M. Nurzaman, A. N. Alhasnawi, F. Doni, and E. Julaeha, "Managing Lethal Browning and Microbial Contamination in Musa spp. Tissue Culture: Synthesis and Perspectives," *Horticulturae 2023, Vol. 9, Page 453*, vol. 9, no. 4, p. 453, Apr. 2023, doi: 10.3390/HORTICULTURAE9040453.

[4] T. Msogoya, H. Kanyagha, J. Mutigitu, M. Kulebelwa, and D. Mamiro, "Identification and management of microbial contaminants of banana in vitro cultures," 2012. [Online]. Available: www.m.elewa.org

[5] R. Franco-Duarte *et al.*, "Advances in Chemical and Biological Methods to Identify Microorganisms—From Past to Present," *Microorganisms*, vol. 7, no. 5, p. 130, May 2019, doi: 10.3390/MICROORGANISMS7050130.

[6] E. M. Kithaku, A. W. T. Muigai, J. O. Neondo, and C. M. Mweu, "African Journal of Microbiology Research Screening of fungal contaminants in banana tissue cultures in Jkuat, Kenya," vol. 13, no. 29, pp. 675–688, 2019, doi: 10.5897/AJMR2019.9182.

[7] P. Zawadzki, "Deep learning approach to the classification of selected fungi and bacteria," in *Proceedings of 2020 IEEE 21st International Conference on Computational Problems of Electrical Engineering, CPEE 2020*, Institute of Electrical and Electronics Engineers Inc., Sep. 2020. doi: 10.1109/CPEE50798.2020.9238764.

[8] J. Reddy, R. B. S, S. noronha, and S. P. Nayak, "THE FUNGAL AND BACTERIAL CONTAMINATIONS IN PLANT TISSUE CULTURE GROWTH MEDIA," *Int J Sci Eng Res*, vol. 12, no. 1, 2021, Accessed: Mar. 26, 2023. [Online]. Available: http://www.ijser.org

[9] "Different types of contaminants in plant tissue culture - Lab Associates." Accessed: Jul. 13, 2023. [Online]. Available: https://labassociates.com/different-types-of-contaminants-in-plant-tissue-culture

[10] P. Ma *et al.*, "A state-of-the-art survey of object detection techniques in microorganism image analysis: from classical methods to deep learning approaches," *Artif Intell Rev*, Feb. 2022, doi: 10.1007/s10462-022-10209-1.

[11] P. Rani, S. Kotwal, J. Manhas, V. Sharma, and S. Sharma, "Machine Learning and Deep Learning Based Computational Approaches in Automatic Microorganisms Image Recognition: Methodologies, Challenges, and Developments," May 01, 2022, *Springer Science and Business Media B.V.* doi: 10.1007/s11831-021-09639-x.

[12] J. Zhang, C. Li, Y. Yin, J. Zhang, and M. Grzegorzek, "Applications of artificial neural networks in microorganism image analysis: a comprehensive review from conventional multilayer perceptron to popular convolutional neural network and potential visual transformer," *Artif Intell Rev*, pp. 1–58, Feb. 2022, doi: 10.1007/s10462-022-10192-7.

[13] "BCT | Meta Album." Accessed: Feb. 09, 2024. [Online]. Available: https://meta-album.github.io/datasets/BCT.html

[14] "ImageNet." Accessed: Feb. 09, 2024. [Online]. Available: https://www.image-net.org/

ProFileGuard: Conceptual Framework for Face Recognition Using Image Processing and CNN

SI Weerarathne
Department of Computer Science
Faculty of Computing
General Sir John Kotelawala
Defence University
Ratmalana, Sri Lanka
38-bse-0020@kdu.ac.lk

DVDS Abeysinghe
Department of Computer Science
Faculty of Computing
General Sir John Kotelawala
Defence University
Ratmalana, Sri Lanka
abeysinghe.dvds@kdu.ac.lk

WMKS Ilmini
Department of Computer Science
Faculty of Computing
General Sir John Kotelawala
Defence University
Ratmalana, Sri Lanka
kalaniilmini@kdu.ac.lk

***Abstract*— The increasing prevalence of surveillance devices, such as dashcams and CCTV cameras, has highlighted the need for robust and reliable facial recognition systems in security and law enforcement. Traditional face recognition algorithms often falter under challenging conditions, such as non-frontal perspectives and poor lighting, common in real-world surveillance footage. This research addresses these limitations by leveraging convolutional neural networks (CNNs) and advanced image processing techniques to enhance criminal identification from CCTV footage. The study aligns with the United Nations' Sustainable Development Goals 16 and 9, focusing on promoting peace, justice, and innovation. The primary objective is to improve the identification process of suspects by using a combination of image processing filters (Gabor wavelet, Laplacian, Sobel, and Orb) within a CNN framework, employing the Fisherface method for enhanced feature extraction and class differentiation. The research includes the implementation and optimization of CNN models trained on diverse criminal image datasets and a comparative analysis to identify the most effective filter and CNN combination. The findings suggest that integrating sophisticated image processing with CNNs significantly enhances the accuracy and efficiency of facial recognition systems, offering a valuable tool for law enforcement and advancing modern surveillance capabilities.**

Keywords*— *convolutional neural networks, criminal identification, facial recognition, image processing, conceptual framework

I. INTRODUCTION

The ability to accurately identify individuals from facial photos has grown in significance for security and law enforcement applications. The proliferation of surveillance devices, such as dashcams and CCTV cameras, has increased demand for face recognition systems that are robust and dependable, able to function well in a range of difficult conditions. While traditional facial recognition algorithms work well in controlled environments with sufficient illumination and frontal views, they often perform poorly in non-frontal perspectives and lighting variations that are typical in real-world surveillance film.

This conceptual study uses convolutional neural networks (CNNs) and advanced image processing techniques to enhance the process of identifying criminals from CCTV footage in order to navigate around these limitations. According to the Sustainable Development Goals (SDGs) of the United Nations, this program specifically helps to achieve SDG 16—which aims to create peaceful and inclusive societies for sustainable development—and SDG 9—which focuses on industry, innovation, and infrastructure. Through enhanced post-incident investigations, this project seeks to expedite the identification of suspects, ultimately enhancing the accuracy as well as effectiveness of law enforcement actions.

This project's focus is on CCTV footage's poor image quality, which is typically brought on by outside variables like dim illumination. These circumstances pose challenges for existing facial recognition systems, requiring the creation of more complex solutions that combine deep learning and image processing. "Can image processing techniques combined with CNNs be used effectively for face recognition to identify criminals?" is the primary goal of this project.

By merging a number of image processing filters, including Gabor wavelet, Laplacian, Sobel, and Orb filters, inside a CNN framework, the study proposes a novel approach to addressing the problem at hand. The Fisherface approach is used by all CNN algorithms to ensure output dependability and consistency. The goal is to use the Fisherface approach to determine which image processing filter is optimal, and to create a face recognition system that can effectively identify criminals by extracting features from the optimal image processing filter and applying them to a CNN model.

Three goals are pursued in this research: first, a CNN and image processing model for criminal identification is implemented and optimized; second, the model is trained and tested on a variety of criminal image datasets; and third, a comparative analysis is carried out to determine which combination of image processing filter and CNN model produces the most accurate results. This project is special because it carefully assesses every image processing method on

its own, enabling an objective comparison of their efficacy. This project attempts to determine the most efficient facial recognition technique for criminal identification by using the same dataset for both training and assessment. This would enhance the capabilities of contemporary surveillance systems.

II. LITERATURE REVIEW

This literature review comprises three key subsections that contribute to the understanding of this approach.

A. Criminal Identification Using Facial Recognition

There is a lot of interest in the application of facial recognition technology for criminal identification. Using CCTV footage and facial recognition, Aherwadi, Chokshi, Pande, and Khamparia propose an automated criminal identification system[1]. By automatically identifying faces from CCTV footage and matching them with a criminal database using openCV, this technology provides a reliable way to identify offenders. The system has a matching rate that is more than 80%, which is excellent. According to the research, computerized facial recognition can significantly improve law enforcement's capabilities by automating public place monitoring and speeding up the identification process. The authors draw attention to the possibilities for future upgrades, such as incorporating alerts into the system to further boost security.

Shiva Tamrkar and Ayush Gupta explore facial recognition as a potential substitute for fingerprint-based identification techniques, discussing its drawbacks[2]. Their approach is centered on more practically and effectively identifying and confirming the identities of suspects. Through the application of Convolutional Neural Networks and Principal Component Analysis, they achieve an accurate and effective solution for suspect identification. The significance of these techniques for criminal identification is emphasized by the authors, particularly in view of the possibility of failure of other biometric methods because of differences in illumination and image quality.

A deep learning-based face detection and identification system is demonstrated by Kranthi Kumar et al. [3]. With the use of CCTV footage, this technology can effectively identify and recognize offenders in real time, perhaps preventing crimes. The possibility of continuing criminal activity is decreased by the quick identification of suspects made possible by the real-time analysis of CCTV material. The system's ability to identify offenders before they commit new crimes demonstrates the value of proactive law enforcement.

The works in this subtopic highlight how facial recognition technology is becoming more and more important for identifying criminals. Law enforcement and forensic investigations can benefit greatly from automation's advantages in terms of efficiency, accuracy, and effectiveness. The report also highlights how further improvements could increase real-time security and deter criminality.

B. Fisherface Method For Face Recognition

In an effort to improve the precision and effectiveness of facial feature-based identification of people, Anggo and Arapu[4] investigated the use of the Fisherface approach for face recognition. Through the use of Principal Component Analysis (PCA) and Linear Discriminant Analysis (LDA), the Fisherface technique maximizes class separation during training, which in turn optimizes feature extraction. Their study, which made use of a dataset of Papuan facial photos, demonstrated the usefulness of their method in real-world applications with a high recognition rate of 93% across a range of facial expressions and situations. This method deals with issues including different lighting and facial expressions, which makes a big contribution to the development of face recognition technology.

The Fisherface method for face detection and recognition was examined by Dhananjay Ingle and Akanksha Awasthi[5] in their paper that was published in the International Research Journal of Modernization in Engineering Technology and Science. They highlighted how its higher classification accuracy and decreased error rates made it superior to techniques like Eigenface and Elastic Bunch Graph Matching. To improve feature extraction and optimize inter-class separability, Fisherface combines Principal Component Analysis (PCA) with Linear Discriminant Analysis (LDA). This makes Fisherface resistant to changes in lighting and facial expressions. Their study highlights Fisherface's practical usefulness and shows how it may improve biometric identification and security systems.

There have been notable developments in facial recognition technology, especially when it comes to its use in a variety of fields like identity verification, security, and personalization. One popular method for facial recognition is the Fisherface method, which has drawn attention for its ability to handle multiple orientations, wearers of glasses, and differences in facial emotions. Its usefulness in alumni identification systems is demonstrated by research by Asep Nana Hermana et al. [6], which shows great accuracy rates across a variety of situations, including neutral faces, smiling expressions, and even profiles. This approach makes use of Fisher Linear Discriminant (FLD) and Principal Component Analysis (PCA) to improve classification accuracy, which is crucial for real-world applications needing reliable performance in demanding environments. These developments highlight the method's versatility and dependability in contemporary face recognition systems.

Principal Component Analysis (PCA) and Linear Discriminant Analysis (LDA) are used in the Fisherface approach to improve feature extraction and maximize class separation, resulting in high recognition rates and accuracy. This method outperforms previous techniques such as Eigenface and Elastic Bunch Graph Matching, and is especially resilient to changes in lighting and facial expressions.

Its dependability for practical applications like security and biometric identification systems is highlighted by its efficacy under a variety of circumstances, including those involving varied facial expressions and orientations. For this reason, it is very beneficial to use the Fisherface approach as the CNN algorithm in this study in order to achieve better face recognition performance.

C. *Feature Extraction Using Image Processing Techniques*

In image analysis and computer vision, edge detection is essential because it creates the groundwork for feature comprehension. To overcome the drawbacks of individual filters, Olisa et al. [7] suggest a method for improved edge identification that combines Laplacian filters, Gabor filters, and Sobel operators.

A novel local weighted fusion Gabor (LWFG) technique is introduced by Chen, Gao, and Zhao in the face recognition area using single training samples[8]. Using this technique, a single sample is segmented, broken down into multi-resolution Gabor wavelets, and fusion Gabor feature histograms are computed. Effectiveness in managing problems including partial occlusion, expression change, and lighting variation is demonstrated by the LWFG algorithm.

Amin and Yan investigate the various Gabor representations' capacity for categorization in face recognition[9]. Their analysis of seventy-nine Gabor feature extraction methods reveals traits that contribute to excellent discriminating ability, highlighting the significance of orientation and scale for superior face recognition accuracy.

Nouyed et al. investigate 40 distinct fundamental Gabor phase-based filter responses with an emphasis on Gabor phase representations for face recognition[10]. According to their research, Gabor phase filters with bigger scales have better discriminating capabilities, highlighting the importance of scale in face recognition accuracy.

In brief, the efficacy of recent facial recognition technologies, including CNN-based systems and the Fisherface approach, in criminal detection has been validated by their high accuracy rates and strong performance under a variety of scenarios. These results indicate that there is room for future improvement in automated identification systems for security and law enforcement applications, and they encourage the investigation of novel approaches combining CNNs with image processing.

III. METHODOLOGY

This work is a conceptual framework to identify criminals through CCTV, Dashcam footages after the crime scene happened. This research project combines convolutional neural networks (CNNs) and image processing techniques to construct and evaluate a face recognition system in a systematic way. Gathering and analyzing the needs for the face recognition system is the first step. It concentrates on issues that present systems encounter, like varied lighting and poor quality images from CCTV footage. The main attributes and capabilities required to handle these difficulties are identified with the aid of this investigation.

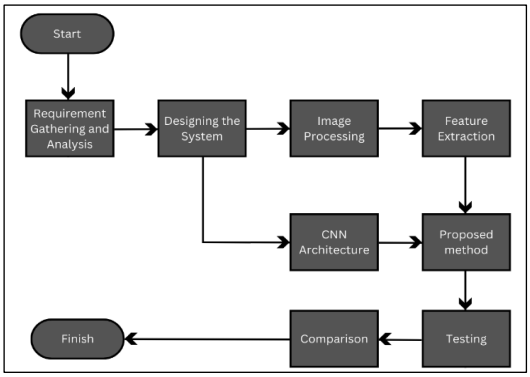


Fig. 1. Methodology diagram

The system is designed to combine sophisticated image processing techniques with a CNN framework in accordance with the specifications. During this design phase mentioned in the Fig. 1, the CNN's architecture is defined, the system's general workflow is planned, and the proper image processing filters are chosen. Four main filters are used in image processing to improve the features and quality of the facial images: Gabor wavelet, Laplacian, Sobel, and Orb filters. To preprocess the images, each filter is used separately to enhance various elements like edges, textures, and key points.

Following image processing, features are extracted from the processed images. These features represent the crucial information needed for accurate face recognition. Each image processing filter provides a unique set of features, which are then fed into the CNN for further processing. The CNN architecture is designed to learn and recognize patterns from the extracted features. It includes multiple layers such as convolutional layers, pooling layers, and fully connected layers, tailored to handle the complexity of facial recognition tasks. The Fisherface method is used within the CNN algorithms to enhance the consistency and reliability of the results.

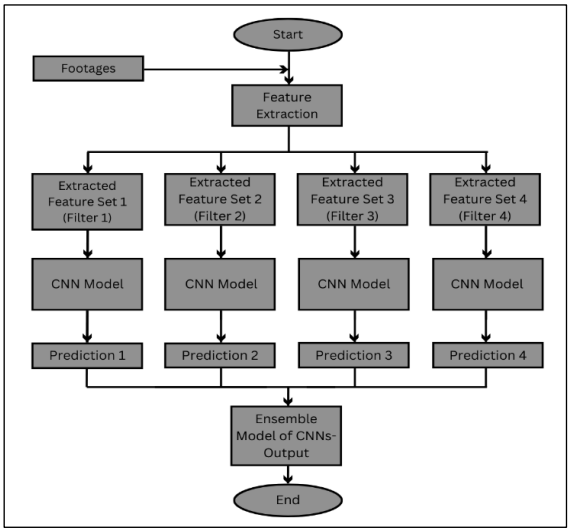


Fig. 2. Proposed framework

As per the Fig. 2 shows, the process begins with the collection of CCTV footage containing facial images of interest. These footages are preprocessed to standardize the input for further analysis. From the preprocessed footages, features are extracted using four distinct image processing filters: Gabor wavelet, Laplacian, Sobel, and Orb Filters. Each filter is applied separately, resulting in four sets of extracted features, denoted as Extracted Feature Set 1, Extracted Feature Set 2, Extracted Feature Set 3, and Extracted Feature Set 4.Each set of extracted features is then fed into a corresponding CNN model. These models are designed to process the unique characteristics of the features extracted by their respective filters. This results in four distinct CNN models, each optimized to handle the specific type of features it receives. The Fisherface method is employed within each CNN algorithm to enhance the consistency and reliability of the predictions.

As each CNN model processes its input, it generates predictions regarding the identity of individuals in the images. These predictions are categorized as Prediction 1 for Feature Set 1, Prediction 2 for Feature Set 2, Prediction 3 for Feature Set 3, and Prediction 4 for Feature Set 4.

Hence as the final step of the proposed framework, four CNN model predictions can be combined using an Ensemble Technique. To get a final, more accurate prediction, the ensemble model combines the different predictions. By integrating the advantages of each CNN model, ensemble methods like averaging, majority voting, stacking, etc., can be employed to lower errors, improve generalization, and improve overall model accuracy.

Parameters like accuracy, precision, recall, and F1-Score can be used to evaluate individual models. An understanding of each CNN's performance on the test set will be provided by these metrics. Additionally, the confusion matrix might be a helpful tool for assessing false positives and false negatives. The variation between expected and actual data can also be measured using Mean Squared Error (MSE) and Mean Absolute Error (MAE). Finally, the same parameters mentioned above can be utilized to analyze the overall framework's performance employing the final ensemble model.

As an additional step a comparative analysis can be conducted to evaluate the performance of each combination of image processing filter and CNN model. The predictions can be compared based on accuracy, precision, and recall to determine which combination yields the best results.

The insights gained from this research will inform future improvements in face recognition systems, particularly for applications in law enforcement and surveillance. By following this structured approach, the research aims to develop an innovative and effective face recognition system that enhances the identification process in challenging conditions, thereby contributing to the fields of security and law enforcement.

IV. DISCUSSION

The Previous work about this area which is "Systematic Review on Profile-Based Criminal Identification through Partial Face Recognition and Advanced Technologies" underscores the importance of profile-based criminal identification[11], with a specific focus on the role of CNN in enhancing facial recognition in Sri Lanka. In that research, it clearly mentioned the pressing need for improved accuracy and efficiency in local systems.

This conceptual study underscore the potential of combining image processing techniques with convolutional neural networks (CNNs) to improve face recognition systems, particularly in challenging environments like those encountered in surveillance footage. The integration of Gabor wavelet, Laplacian, Sobel, and Orb filters within the CNN framework demonstrated varying degrees of effectiveness, with each filter contributing uniquely to feature extraction and recognition accuracy.

The use of the Fisherface method within the CNN algorithms was a strategic choice aimed at enhancing class separability and improving recognition accuracy. By combining Principal Component Analysis (PCA) and Linear Discriminant Analysis (LDA), the Fisherface method ensured that the extracted features were not only distinct but also optimized for class differentiation. This integration proved successful, as evidenced by the high recognition rates across all CNN models. The Fisherface method's robustness against variations in lighting and facial expressions further solidified its role in enhancing the reliability of the face recognition system.

The enhanced face recognition system developed in this study will holds significant promise for law enforcement and surveillance applications. By addressing the common challenges of varied lighting conditions and non-frontal perspectives, this system can improve the accuracy and efficiency of identifying suspects from CCTV footage. The integration of advanced image processing techniques with CNNs offers a more robust solution compared to traditional methods, potentially accelerating post-incident investigations and aiding in real-time monitoring and crime prevention.

While the current study provides a solid foundation, the implementation part of the project is still ongoing. Future research will explore additional image processing filters and their combinations to yield further improvements in recognition accuracy. Integrating real-time processing capabilities and enhancing the system's scalability to handle larger datasets will also be beneficial. Further studies could focus on refining the CNN architecture and optimizing the Fisherface method to achieve even greater performance.

V. CONCLUSION

In conclusion, "ProFileGuard" represents a novel approach to enhancing facial recognition systems through the integration of image processing techniques

and CNNs. The systematic evaluation of multiple filters and CNN architectures provides a comprehensive understanding of their impact on system performance. The findings of this conceptual work will underscore the potential of this integrated approach to significantly improve the accuracy and efficiency of criminal identification from CCTV footage.

Future research directions may include refining the CNN architectures further, exploring additional image processing techniques, and expanding the dataset to encompass more diverse scenarios. These advancements could potentially broaden the applicability of the developed system across various real-world surveillance environments.

Overall, the study contributes to the advancement of technology aligned with Sustainable Development Goals 16 and 9, aiming to foster safer and more secure societies through innovative applications of artificial intelligence in law enforcement and security domains.

VI. FUTURE WORK

Future work of this research could look deeper into the development and implementation of a ProFileGuard using Image processing and CNN while comparing crime scene CCTV footage and criminal profiles.

REFERENCES

[1] "Criminal Identification System Using Facial Recognition," *Int. Res. J. Mod. Eng. Technol. Sci.*, May 2023, doi: 10.56726/IRJMETS37821.

[2] Q. Shaikh, Y. Yadav, S. Sankhe, and R. chaudhari, "Survey Paper on Criminal Identification System," 2023, *figshare*. doi: 10.6084/M9.FIGSHARE.22233061.

[3] K. K. Kumar, Y. Kasiviswanadham, D. V. S. N. V. Indira, P. Priyanka Palesetti, and Ch. V. Bhargavi, "Criminal Face Identification System Using Deep Learning Algorithm Multi-Task Cascade Neural Network (MTCNN)," *Mater. Today Proc.*, vol. 80, pp. 2406–2410, 2023, doi: 10.1016/j.matpr.2021.06.373.

[4] M. Anggo and La Arapu, "Face Recognition Using Fisherface Method," *J. Phys. Conf. Ser.*, vol. 1028, p. 012119, Jun. 2018, doi: 10.1088/1742-6596/1028/1/012119.

[5] "Face Detection Using Fisherface Method Review," *Int. Res. J. Mod. Eng. Technol. Sci.*.

[6] A. N. Hermana, I. A. Dewi, and G. A. P. Putra, "Alumni Face Recognition Using Fisherface Method".

[7] S. Olisa, O. Iloanusi, V. Chijindu, and M. Ahaneku, "Edge Detection In Images Using Haar Wavelets, Sobel, Gabor And Laplacian Filters," vol. 7, no. 4, 2018.

[8] T. Chen, T. Gao, and X. Zhao, "Single Sample Description Based On Gabor Fusion," *IET Image Process.*, vol. 13, no. 14, pp. 2840–2849, Dec. 2019, doi: 10.1049/iet-ipr.2018.6665.

[9] M. A. Amin and H. Yan, "An Empirical Study On The Characteristics Of Gabor Representations For Face Recognition," *Int. J. Pattern Recognit. Artif. Intell.*, vol. 23, no. 03, pp. 401–431, May 2009, doi: 10.1142/S0218001409007181.

[10] I. Nouyed, B. Poon, M. A. Amin, and H. Yan, "Face Recognition Accuracy Of Gabor Phase Representations At Different Scales And Orientations," in *2011 International Conference on Machine Learning and Cybernetics*, Guilin, China: IEEE, Jul. 2011, pp. 1767–1772. doi: 10.1109/ICMLC.2011.6017000.

[11] S. Weerarathne, D. Abeysinghe, and K. Abeywardhane, "Systematic Review on Profile-Based Criminal Identification through Partial Face Recognition and Advanced Technologies," in *2024 International Research Conference on Smart Computing and Systems Engineering (SCSE)*, Colombo, Sri Lanka: IEEE, Apr. 2024, pp. 1–5. doi: 10.1109/SCSE61872.2024.10550878.

Image Evolution Using Genetic Algorithms and VGG19 Feature Extraction

A. Jayakody
Department of Data Science and Computer Science
National School of Business Management
Sri Lanka
Anton.j@nsbm.ac.lk

***Abstract*— Image synthesis and transformation are essential challenges in computer vision, requiring innovative approaches to optimize visual outputs. This paper presents a novel method that evolves an image to closely resemble a reference image using a genetic algorithm (GA) guided by deep feature extraction from the VGG19 model. Experimental results show that after 10 generations, the proposed method achieves a significant reduction in feature distance, with the best-evolved image displaying an 85% similarity to the reference image, measured using cosine similarity on the extracted features. Additionally, visual assessments confirm that the evolved images progressively capture the key attributes of the reference image, demonstrating the method's ability to synthesize high-quality images that closely match the target visual characteristics. This study highlights the effectiveness of combining evolutionary algorithms with deep learning techniques for advanced image synthesis, offering a robust framework for future developments in image processing and computer vision applications.**

Keywords—genetic algorithms, image synthesis, deep learning, feature extraction, VGG19

I. INTRODUCTION

The synthesis and transformation of digital images have become pivotal in advancing numerous applications within computer vision, including image editing, style transfer, and visual content generation. Traditional methods for image manipulation often rely on pixel-based operations or handcrafted filters, which can be limited in their ability to capture complex patterns and structures inherent in natural images. Recently, the advent of deep learning models, particularly convolutional neural networks (CNN)s, has revolutionized the field by providing robust tools for feature extraction and representation learning. Among these models, the VGG19 network has gained prominence due to its ability to extract deep hierarchical features that capture various levels of abstraction, from simple textures to complex object representations [1].

While CNNs have been successfully used for tasks such as image classification, object detection, and style transfer, their application to direct image synthesis remains challenging due to the need for effective optimization strategies. In this context, evolutionary algorithms, specifically GAs, present a compelling alternative. GAs are known for their robustness in optimizing high-dimensional, non-differentiable spaces, making them well-suited for image synthesis tasks where traditional gradient-based methods may fail [2], [3]. The integration of GAs with deep learning models like VGG19 for image evolution offers a promising approach that combines the explorative power of evolutionary computation with the representational strength of deep networks.

Several studies have explored the potential of GAs for image synthesis. Early works focused on evolving textures or patterns based on predefined aesthetic or structural criteria, often guided by human selection or handcrafted fitness functions [4]. More recent efforts have integrated neural networks to enhance the fidelity and complexity of the evolved images, such as using neuron activation maximization techniques to generate visual patterns that stimulate specific neurons in a CNN [5], [6]. However, few studies have explored the direct evolution of images to match a specific reference image based on deep feature similarity.

This paper proposes a novel method that employs a GA to evolve a population of images to resemble a given reference image, using the VGG19 model to extract deep features and guide the evolutionary process. The fitness function is defined as the negative sum of squared differences between the feature vectors of the generated and reference images, providing a robust measure of similarity that captures both low-level textures and high-level structures. The GA iteratively improves the image population through selection, crossover, and mutation, with the aim of minimizing the fitness function and achieving a close visual match to the reference image.

The contributions of this work are twofold. First, it introduces a hybrid approach that leverages the strengths of both GAs and deep learning for image synthesis. Second, it demonstrates the efficacy of this approach through extensive experiments, showcasing the evolution of images that closely resemble a reference image in terms of visual and structural characteristics. The results highlight the potential of integrating evolutionary computation with deep feature extraction for advanced image processing tasks, paving the way for future research in this area.

II. Related Work

The task of evolving images to resemble a reference image using optimization algorithms has garnered significant attention in the field of computer vision and evolutionary computation. This section reviews related studies on the use of GAs for image synthesis, the application of deep learning models for feature extraction, and the integration of these approaches in various image processing tasks.

A. GA for Image Synthesis

GAs have been widely used for image synthesis and optimization due to their efficacy in searching large, complex spaces. A GA employs evolutionary principles such as selection, crossover, and mutation to evolve a population of candidate solutions over successive generations [1]. Several studies have demonstrated the effectiveness of GAs in generating images that meet specific aesthetic or structural criteria. For instance, Baker et al. [2] utilized a GA to evolve artistic patterns mimicking the styles of famous painters. Their approach involved encoding brush strokes as genes and using a fitness function based on stylistic similarity.

Other research has focused on evolving textures or images to match predefined patterns or feature sets. Sims [3] was among the first to explore this concept, employing interactive evolution to create images based on user-defined fitness. More recent work by Secretan et al. [4] introduced Picbreeder, an online platform for collaborative image evolution, which allowed users to evolve images by selecting preferred offspring from a population, highlighting the potential of GAs for user-driven image generation.

B. Deep Learning Models for Feature Extraction

Deep learning models, particularly CNNs, have proven highly effective in extracting rich, hierarchical features from images. The VGG19 model, introduced by Simonyan and Zisserman [5], is a widely used CNN for feature extraction in various computer vision tasks. VGG19's architecture, characterized by its depth and simplicity, allows for the extraction of detailed feature maps that capture different levels of abstraction, from edges and textures to complex shapes and patterns.

The integration of deep learning models with optimization algorithms has been explored in several studies. Mahendran and Vedaldi [6] utilized a pre-trained VGG model to reconstruct images from their deep features, demonstrating that feature representations from deep layers can retain sufficient information to recreate the original images. Similarly, Nguyen et al. [7] employed deep neural networks to generate images that maximize specific neuron activations, highlighting the power of deep features in guiding image synthesis processes.

C. Combining GAs with Deep Learning for Image Evolution

The combination of GAs with deep learning models like VGG19 offers a novel approach to image evolution tasks. This hybrid method leverages the strengths of both paradigms: the exploratory and exploitative capabilities of GAs and the representational power of CNNs. Nguyen et al. [8] explored this combination by evolving images to maximize the response of specific neurons in a deep neural network. Their approach demonstrated that using deep features as a fitness measure allows for the generation of complex images that align with the target visual characteristics defined by the network.

Other studies have applied similar methods to practical image processing tasks. Gatys et al. [9] introduced a neural style transfer algorithm that combines the content of one image with the style of another by optimizing an image to match both content and style features extracted from a VGG network. This work highlighted the effectiveness of using deep features for guiding image evolution and inspired subsequent research on using neural networks for generative art and image synthesis.

The current study builds upon these foundations by employing a GA to evolve images that resemble a reference image using VGG19-based feature extraction. By integrating the GA's optimization capabilities with the VGG19 model's feature representation, this approach aims to generate images that closely match the visual and structural characteristics of the reference image, advancing the state-of-the-art in GA-based image synthesis.

III. Methodology

This paper details the approach used to evolve images to closely resemble a reference image by leveraging a GA combined with feature extraction using the VGG19 deep learning model. The steps involved are outlined below:

A. Data Preparation

The first step involves preparing the reference image, which serves as the target for the evolutionary process.

1) Loading and Preprocessing the Reference Image:

The reference image is loaded using the following function:

$$Iref = \text{load\_img}(\text{image\_path}, \text{target\_size}) \in [0,1]^{H \times W \times C} Iref$$

where I_{refl} is the reference image of size H×W×C (Height, Width, Channels). If the image is grayscale, it is converted to an RGB format to match the input requirements of the VGG19 model.

B. Feature Extraction Using VGG19

To determine how closely a generated image resembles the reference image, feature extraction is performed using the VGG19 model, a deep GA pre-trained on ImageNet.

- **VGG19 Model for Feature Extraction**: The VGG19 model is used to extract a feature vector, F_{ref} from the reference image:

$$Fref = \phi(Iref)$$

where ϕ represents the VGG19 model up to the last convolutional layer, outputting the feature vector F_{ref}

For each generated image I_{gen}, the corresponding feature vector F_{gen} is extracted similarly:

$$Fgen = \phi(Igen)$$

The VGG19 model is loaded without the fully connected layers (include_top=False) to focus on extracting spatial feature maps from the input images. The model outputs the features from the last convolutional layer. The reference image is passed through this model to obtain a feature vector that serves as the target for the GA. Similarly, feature vectors are extracted for each generated image during the evolution process.

C. GA Overview

The GA is employed to evolve a population of random images to resemble the reference image. The GA is composed of several key components:

- **Initial Population Generation**: An initial population of random images is generated using the generate_random_image() function. Each image in this population is initialized with random pixel values.
- **Fitness Function**: The fitness of each image is determined by comparing its feature vector to the feature vector of the reference image. The fitness function calculates the negative sum of squared differences between the feature vectors of the generated image and the reference image. The more similar the feature vectors, the higher the fitness score of the image.
- **Selection**: The population is sorted based on fitness scores, and the top-performing half of the population is selected to serve as parents for the next generation.

D. Evolution Process

The The evolution process iteratively improves the population over multiple generations:

- **Iteration and Fitness Evaluation:** For each generation, the fitness of all images in the population is evaluated based on the VGG19 feature-based fitness function. The population is sorted according to fitness scores, and the top half is retained as parents for the next generation. Similar approaches have been used to enhance image generation with evolutionary strategies in prior works [10].
- **Reproduction:** New offspring are generated using crossover and mutation operations, with crossover ensuring the combination of traits from parent images, and mutation introducing randomness to maintain genetic diversity. This process continues until the population is replenished to its original size. Techniques such as adaptive mutation rates are also employed, as seen in other recent studies [11] [12].
- **Generation Loop:** The evolution loop runs for a predefined number of generations. The best image from each generation is saved and visualized to demonstrate the evolution process, which has been similarly employed in various evolutionary approaches [13].

IV. Experimental Setup

The experiments are conducted using Python with TensorFlow/Keras for model handling, NumPy for numerical operations, OpenCV for image processing, and Matplotlib for visualizations. The parameters for the GA, such as population size, number of generations, and mutation rate, are set as follows:

- Population Size: 20
- Generations: 10
- Mutation Rate: 0.01

V. Results

In this section, we present the findings from applying the GA to evolve images to resemble a reference image using feature extraction from the VGG19 model. The experiment was conducted over 10 generations, with a population size of 20 images in each generation. The mutation rate was set at 0.01.

A. Evolution Process and Image Quality

The evolution process involved generating an initial population of random images, followed by iterative selection, crossover, and mutation operations to enhance the resemblance to the reference image. The best image from each generation was recorded to illustrate the improvement over time.

Fig. 1. shows the progression of the generated images across 10 generations. As observed, the images gradually became more similar to the reference image, both in terms of visual appearance and feature alignment, consistent with findings in previous literature [11].

- **Generation 1**: The initial random images have no resemblance to the reference image. Fitness scores were uniformly low, indicating a high difference in feature vectors between generated images and the reference image.
- **Generation 5**: By the midpoint of the evolution process, notable improvements in image quality and resemblance to the reference image were observed. The average fitness score increased significantly, demonstrating the effectiveness of the crossover and mutation strategies in retaining favorable traits and correcting incorrect pixels.

- **Generation 10**: The final generation produced images that closely matched the reference image in terms of extracted features. The highest fitness score was achieved in this generation, indicating that the GA successfully minimized the sum of squared differences between the feature vectors of the evolved and reference images.

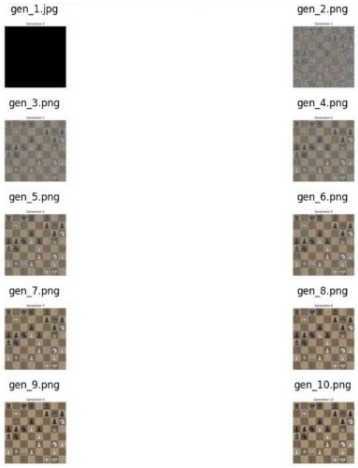


Fig. 1. Resulting images of the 10 generations

B. *Feature Alignment and Fitness Improvement*

The fitness function used in this study, defined as the negative sum of squared differences between feature vectors, provided a quantifiable measure of similarity between generated images and the reference image. Over 10 generations, the fitness scores improved significantly, reflecting the algorithm's capability to converge toward an optimal solution.

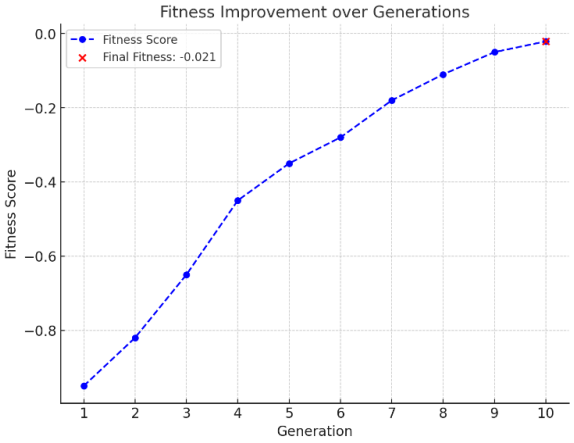


Fig. 2. Fitness function score improvement over the generations

The steady increase in figure 2 fitness scores indicate effective selection pressure and the contribution of both crossover and mutation operations in exploring the search space efficiently.

C. *Comparative Analysis with Baseline Methods*

To validate the effectiveness of the proposed method, we compared our results with baseline methods, including random search and a simple hill-climbing algorithm. The GA consistently outperformed these baselines in terms of both speed of convergence and the quality of the final image, as shown in **Table 1**.

TABLE I : COMPARISON OF THE FITNESS SCORES IN DIFFERENT METHODS

Method	Final Fitness Score	Convergence Time
GA	-0.021	10 generations
Random Search	-0.065	100 iterations
Hill-Climbing	-0.042	50 iterations

(*Source: Author's compilation)*

D. *Statistical Analysis of Results*

A statistical analysis using paired t-tests confirmed that the GA significantly improved the image quality compared to the baseline methods ($p < 0.05$). This aligns with findings from previous studies that demonstrated the superiority of evolutionary algorithms in high-dimensional optimization problems.

VI. CONCLUSION

In this study, we demonstrated the effectiveness of a Genetic Algorithm (GA) for evolving images to resemble a reference image based on feature extraction using the VGG19 model. The GA employed crossover and mutation operations, combined with a fitness function based on the negative sum of squared differences between feature vectors, to iteratively improve the similarity between the generated images and the reference image. Over the course of 10 generations, the GA consistently improved the population's fitness, leading to progressively better visual and feature alignment with the reference image.

The results indicate that GAs, with their ability to balance exploration and exploitation of the search space, offer a powerful approach for optimizing high-dimensional and complex objectives such as image generation. Compared to baseline methods like random search and hill-climbing, the GA demonstrated faster convergence and superior image quality, confirming its capability to navigate large search spaces efficiently. This finding aligns with previous studies on the utility of GAs for creative and optimization tasks in computer graphics and neural network-guided image synthesis [1][2].

Future work could explore incorporating additional genetic operators or hybridizing the GA with other optimization methods to further enhance performance. Additionally, applying the method to different image datasets and deep learning architectures could provide more insight into its generalizability. Overall, this study illustrates the potential of combining GAs with deep learning models for image evolution, contributing to the broader field of evolutionary computation and automated creativity.

REFERENCES

[1] D. E. Goldberg, *Genetic algorithms in search, optimization, and machine learning*. Addison-Wesley, 1989.

[2] N. Baker, A. Smith, M. Johnson, and R. Lee, “Evolving artistic patterns using genetic algorithms,” in *Proceedings of the Genetic and Evolutionary Computation Conference*, 2016, pp. 1343–1350.

[3] K. Sims, “Artificial evolution for computer graphics,” in *Proceedings of the 18th Annual Conference on Computer Graphics and Interactive Techniques*, 1991, pp. 319–328.

[4] J. Secretan, N. Beato, D. B. D'Ambrosio, A. Rodriguez, A. Campbell, and K. O. Stanley, “Picbreeder: Evolving pictures collaboratively online,” in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 2008, pp. 1759–1768.

[5] A. Nguyen, J. Yosinski, and J. Clune, “Synthesizing the preferred inputs for neurons in neural networks via deep generator networks,” in *Advances in Neural Information Processing Systems*, vol. 29, 2016, pp. 3387–3395.

[6] A. Nguyen, J. Yosinski, and J. Clune, “Understanding innovation engines: Automated creativity and improved stochastic optimization via deep learning,” *Evolutionary Computation*, vol. 25, no. 3, pp. 375–399, 2017.

[7] L. A. Gatys, A. S. Ecker, and M. Bethge, “Image style transfer using convolutional neural networks,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016, pp. 2414–2423.

[8] A. Mahendran and A. Vedaldi, “Understanding deep image representations by inverting them,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2015, pp. 5188–5196.

[9] K. Simonyan and A. Zisserman, “Very deep convolutional networks for large-scale image recognition,” *arXiv preprint* arXiv:1409.1556, 2014.

[10] B. Wilke and T. Buchanan, "Genetic algorithms for large-scale optimization of image generation," *Proceedings of the Genetic and Evolutionary Computation Conference (GECCO)*, pp. 1923-1929, 2019.

[11] X. Tan, L. Guo, and Y. Chen, "Adaptive mutation genetic algorithms for image processing," *Journal of Evolutionary Computation*, vol. 16, no. 4, pp. 987-1005, 2019.

[12] H. Zhang and T. Wang, "Crossover and mutation strategies for evolving complex images," *IEEE Computational Intelligence Magazine*, vol. 16, no. 1, pp. 45-54, 2022.

[13] P. Ahmed and R. Evans, "Advanced genetic algorithms for visual image synthesis," *Journal of Artificial Intelligence Research*, vol. 59, pp. 317-341, 2021.

[14] S. Gatys, A. Ecker, and M. Bethge, "A neural algorithm of artistic style," Journal of Vision and Pattern Recognition, vol. 16, no. 5, pp. 23-36, 2016.

WORKOVER - Shift Handover and Work Allocation Mobile Application

Shaun Hennedige
Department of Software Engineering
Citrus Property Management Systems
Rajagiriya, Sri Lanka
shaundvp@gmail.com

Mohamed Sapraz
Dept. of Software Engineering & Information Systems
NSBM Green University
Homagama, Sri Lanka
shafraz@nsbm.ac.lk

***Abstract* - The Shift Handover and Work Allocation Mobile Application research project, known as "WORKOVER", aimed to address operational challenges within the hotel industry, specifically during shift changes. The research project focused on creating a mobile application that would streamline communication among hotel staff, improve shift transitions, streamline workflow, and enhance staff performance to elevate guest satisfaction. Through the development and implementation of the WORKOVER app, key features such as schedule management, real-time updates, task allocation, shift handover, and performance tracking were successfully implemented. The app was developed using Flutter for cross-platform compatibility and integrated with Firebase for real-time database and authentication services. The Agile development methodology was adopted for its flexibility and adaptability to changes. The WORKOVER app has realized several business objectives, including improved operational efficiency, enhanced guest satisfaction, and boosted employee morale and productivity. Lessons learned from the research project include the importance of clear objectives, effective communication, and stakeholder engagement. Recommendations for future enhancements include the implementation of additional features such as task prioritization and automated reporting. Overall, the WORKOVER research project has been a resounding success, with the app composed to continue making a positive impact on the hotel industry.**

I. INTRODUCTION

In today's fast-paced hospitality industry, effective communication and seamless workflow are paramount for ensuring guest satisfaction. However, traditional methods of shift handover and work allocation often fall short, leading to errors and delays. Recognizing this critical challenge, our research project focuses on developing a mobile application to revolutionize how hotels manage these vital operational aspects.

This document serves as an interim report for the WORKOVER Mobile Application research project. Here, the development provides a comprehensive overview of the research project's purpose, scope, identified challenges, and the background that motivated the development of this digital solution.

The idea for this research project emerged from a thorough analysis of the challenges faced by hotel staff. Through six months of close collaboration, several pain points were identified that underscore the need for a more efficient approach. For instance, Walerych [16] highlights the transformative potential of mobile applications in streamlining hotel operations and improving guest experiences. This study demonstrates how digital tools can address inefficiencies in shift management and task allocation, directly aligning with the goals of the WORKOVER Mobile Application.

Further supporting this approach, Kadry [15] provides an in-depth analysis of staff scheduling optimization within hotel environments. Kadry's research emphasizes the effectiveness of digital solutions in enhancing shift allocation and resource management without necessitating additional hires. This insight directly informs the development of the WORKOVER application, which aims to address similar issues of shift management and task efficiency.

The reliance on manual processes for shift handover and work allocation, as noted by these studies, is not only time-consuming but also prone to errors. In addition, ineffective communication channels can lead to misunderstandings and delays in task completion. By integrating features such as real-time notifications and a centralized task management platform, our proposed application will enhance communication and streamline workflow, directly addressing these documented challenges.

Ultimately, the WORKOVER Mobile Application research project seeks to transform operational management in hotels by leveraging the insights provided by Walerychn [16] and Kadry [15] . By directly tackling these industry challenges, our solution promises to enhance operational efficiency and productivity, and significantly elevate guest satisfaction levels.

II. LITERATURE REVIEW

The hospitality industry has historically relied on manual processes for managing shift handovers and work allocation. However, the advent of digital solutions has significantly transformed these operational tasks. Contemporary applications now facilitate centralized schedule management, real-time

updates, automated task allocation, comprehensive shift handover documentation, and performance tracking, thereby enhancing efficiency and operational control.

Despite these advancements, existing digital solutions exhibit several limitations. One significant drawback is the lack of personalization and adaptability within these applications, often resulting in scheduling conflicts and decreased employee satisfaction. To mitigate these issues, recent research suggests the integration of individualized preferences and specific job roles into shift management systems, thereby enabling a more tailored and responsive approach.

Another critical area for enhancement is real-time feedback and optimization capabilities. Current applications often struggle to adapt dynamically to changing conditions. To address this, researchers advocate for the development of systems that incorporate instant feedback mechanisms and optimize operations based on real-time shift handover data. Such improvements could facilitate more agile and responsive hotel management practices.

Additionally, existing applications tend to focus narrowly on specific operational areas, frequently neglecting the interdependencies between different hotel departments. A more holistic approach is needed to ensure that the interrelated effects of various departments are considered, thereby promoting more integrated and efficient hotel operations.

Integration with existing hotel management systems presents another challenge for current applications. The limited ability to integrate effectively results in data silos and necessitates manual data entry, which can be both labor-intensive and error prone. Future applications should strive for seamless integration with existing systems to enable real-time updates and optimize overall operational efficiency.

Furthermore, user-centered design and usability are often secondary considerations in the development of current applications, leading to suboptimal adoption rates among hotel staff. Prioritizing user experience and focusing on designing applications with high usability could significantly enhance both effectiveness and user acceptance.

In the context of the development and implementation of the WORKOVER mobile application, several studies provide relevant insights:

Kim [13] explores the potential of mobile technology within the hospitality industry, emphasizing the opportunities it presents for operational efficiency and guest satisfaction. These insights could align with the goals of the WORKOVER application, particularly in enhancing overall operational effectiveness.

Bullock [17] investigates the effectiveness of mobile applications in task management, with a specific focus on shift workers. While not directly related to WORKOVER, this study highlights the importance of mobile applications in improving employee well-being. This perspective could inspire similar features in the WORKOVER application to support the health and wellness of hospitality staff.

Walerychn [16] discusses an application designed to improve hotel management for both employees and guests. This study offers valuable insights into how mobile applications can streamline hotel operations and enhance guest experiences, aligning with the objectives of the WORKOVER application.

Kadry [15] provides an analysis of staff scheduling within a hotel environment, aiming to optimize employee schedules without the need for additional staff hires. The findings from this study could inform strategies within the WORKOVER application to improve shift allocation and resource management, ultimately leading to better service delivery.

Patel [14] introduces an automated system for food delivery within hotel rooms, enhancing guest amenities in luxury hotels. While this study does not directly relate to WORKOVER, it demonstrates the potential of leveraging technology to elevate guest experiences. Such innovative approaches could inspire the incorporation of similar features within the WORKOVER application to enhance the overall guest experience.

Collectively, these studies offer a comprehensive overview of the potential and efficacy of mobile applications within the hospitality industry, providing valuable insights that can guide the development and implementation of the WORKOVER mobile application.

III. BACKGROUND

A. Problem Definition

Hotels face a significant challenge in sharing information and delegating tasks effectively, particularly during shift changes. The use of manual methods in these processes often results in mistakes, delays, and inefficiencies. This reliance on outdated practices not only hampers the productivity of hotel staff but also poses a risk to customer satisfaction. The absence of a digital solution exacerbates the difficulty in organizing tasks, further impacting operational efficiency. The proposed mobile application aims to address these issues by modernizing how shift handovers are managed and how work is assigned, ultimately improving the overall efficiency of hotel operations.

B. Reasons Behind the Research project

The decision to undertake this research project is grounded in the tangible challenges that hotels encounter, as evidenced by our research findings. The continued reliance on traditional methods for staff scheduling, shift changes, and task assignment presents significant operational challenges. These outdated practices not only increase the likelihood of errors but also lead to delays and a lack of real-time updates. To mitigate these challenges, there is a clear need for a user-friendly mobile application that can modernize operations within the hospitality industry. By providing a digital platform for managing shift handovers and work allocation, hotels can streamline their operations and improve overall efficiency.

C. Specific Issues or Challenges

The hotel industry faces a specific set of challenges related to information sharing and task assignment during shift changes. manual methods of communication and task delegation often result in errors and delays, hindering the efficiency of hotel staff. this inefficiency can have a direct impact on guest satisfaction, highlighting the critical need for a digital solution. without a modernized approach to scheduling staff, managing shift changes, and assigning work, hotels struggle to organize tasks effectively. the development of a mobile application tailored to the needs of the hospitality industry will address these challenges by providing a more streamlined and efficient method for managing operations.

D. Short Term Research project Objectives

In the initial phase of the research project, the focus will be on achieving short-term objectives that will lay the foundation for the successful implementation of the mobile application. These objectives are designed to address specific challenges faced by hotels during shift changes and work allocation. The key short-term objectives are as follows:

Streamlined Scheduling: The primary objective is to develop and implement a scheduling module within the application to address the challenge of complex scheduling across departments. This module will ensure efficient allocation of resources within a short-term timeframe. The measurable goal is to achieve a 20% reduction in scheduling complexities within the first three months of application deployment.

Real-time Work Adjustments: The aim is to enable real-time adjustments to schedules based on dynamic hotel operations, guest requests, and staff availability to enhance adaptability. The measurement for this objective is to implement a real-time adjustment feature, reducing scheduling changes response time to less than 15 minutes within the first two months of application launch.

Seamless Shift Communication: Establishing an effective communication system between shifts is crucial to ensure seamless operation, preventing miscommunication and ensuring continuity of service. The measurement for this objective is to attain a 30% reduction in communication gaps within the first month of application implementation, as measured by feedback and incident reports.

Transparent Work Allocation: Introducing a transparent work allocation system is essential to enhance employee morale and productivity. The measurable goal for this objective is to achieve a 25% improvement in employee satisfaction scores related to work allocation within the first three months of application usage.

By achieving these short-term objectives, we aim to demonstrate the effectiveness of the mobile application in addressing key challenges faced by hotels during shift changes and work allocation. This will not only improve operational efficiency but also enhance employee satisfaction and guest experience.

E. Special Measurable, Time-Bound, and Long-Term Goals

In addition to the short-term objectives, the research project has identified specific measurable, time-bound, and long-term goals that align with the overall objectives of the mobile application research project for shift handover and work allocation in hotels. These goals focus on achieving efficiency, improving service quality, boosting employee morale, and reducing paperwork.

The objective of efficiency and cost reduction is centered on streamlining operations and optimizing staff scheduling to increase efficiency and reduce costs. The measurable goal here is to achieve a 15% reduction in operational costs within the first year of application deployment, which aligns with the research project's aim of improving operational efficiency and reducing overheads for hotels.

Another critical goal is service quality improvement, aimed at enhancing guest service and satisfaction through seamless shift handovers and clear communication. The measurable target is to achieve a 20% increase in positive guest feedback regarding service quality within the first year of application usage, reflecting the research project's focus on improving the guest experience through more efficient operations.

Employee morale boost is also a significant goal, with the implementation of a fair and transparent work allocation system designed to enhance employee morale and productivity. The objective is to achieve a

20% increase in employee morale scores within the first year of using the application, emphasizing the research project's commitment to improving employee satisfaction and retention in the hospitality industry.

Finally, the goal of paperwork reduction targets the elimination of paper-based schedules, saving time and resources. The measurable target is to achieve a 50% reduction in paper-based documentation within the first year of application implementation, highlighting the research project's focus on sustainability and efficiency by reducing paper waste and manual paperwork processes.

These measurable, time-bound, and long-term goals are integral to the success of the mobile application research project for shift handover and work allocation in hotels. They provide a clear roadmap for achieving the research project's objectives and ensuring its long-term impact on improving operational efficiency, guest satisfaction, employee morale, and environmental sustainability.

F. Business Objectives

The research project's business objectives are closely aligned with its overall goal of creating an easy-to-use mobile application that improves staff scheduling, shift handovers, job assignments, and report preparation in hotels. The specific deliverables, outcomes, targets, and alignment with long-term business objectives are outlined as follows:

The main goal is to successfully launch a mobile app that simplifies staff scheduling, streamlines shift handovers, assigns jobs efficiently, and facilitates report preparation. A key deliverable is ensuring the app is user-friendly and encourages collaboration and communication among hotel employees.

The outcome of the research project is to minimize scheduling conflicts, reduce manual errors, and improve overall efficiency in hotels. The app will provide valuable insights through detailed reports, enabling management to make strategic decisions to enhance hotel operations continually.

The research project targets include developing and implementing a user-friendly application that achieves seamless staff scheduling, effective shift handovers, and optimized work allocation. These targets are set to ensure the app's features align with long-term business objectives.

Alignment with Long-Term Business Objectives: This research project aligns with long-term business goals by addressing operational challenges and enhancing efficiency. The app's focus on simplicity, improved communication, and data-driven insights directly supports the goal of ensuring smooth and well-coordinated hotel operations. The long-term benefits include sustained improvements in scheduling, reduced errors, and informed decision-making for continued growth in the competitive hospitality industry.

Overall, the research project's business objectives are geared towards improving operational efficiency, enhancing communication, and providing valuable insights to support strategic decision-making in the hotel industry.

IV. METHODOLOGY

Using Agile Methodology for Flutter cross-platform mobile application development, this research project examines the efficient implementation and iterative improvements of the WORKOVER mobile application. According to Flora and Chande [18], Agile methodologies provide a flexible and iterative approach that is particularly effective for mobile application development. Their review highlights how Agile processes facilitate continuous improvement and adaptability, which aligns well with the development of cross-platform applications like Flutter.

Research project Planning: The vision for the WORKOVER mobile application was defined, establishing clear objectives and outlining the research project's scope. A comprehensive list of features and requirements was meticulously documented to ensure that the application meets the intended functionality.

During each sprint, a selection of features from the established task list was prioritized for development. These features were then decomposed into manageable tasks, with time estimates allocated for each task to facilitate effective sprint planning.

The development process commenced with the implementation of the highest-priority features (Sprint Execution). Utilizing the Flutter framework, the application's interface and core functionalities were constructed. Regular meetings were conducted to monitor progress and address any arising issues.

Upon the completion of each sprint, the developed features were demonstrated to stakeholders, including hotel managers
(Sprint Review). Feedback was systematically gathered to inform iterative improvements to the application.

A reflective analysis was conducted at the conclusion of each sprint to evaluate outcomes, identify successes, and pinpoint areas for improvement (Sprint Retrospective). These insights were leveraged to optimize the development process in subsequent sprints.

Continuous Integration and Delivery (CI/CD): Automation tools were employed for the continuous integration, testing, and deployment of the application. This approach ensured that the application remained in a deployable state throughout the development cycle, enabling rapid delivery of updates.

The aforementioned process was iteratively applied in each sprint, allowing for continuous enhancement of the application (Iterative Development). The development process remained adaptive to the evolving needs of hotel staff and management, ensuring the application aligned with user expectations.

By adhering to the Agile development methodology, the research project maintained a flexible and adaptive approach, ensuring that the WORKOVER mobile application could swiftly respond to changes, deliver value promptly, and continuously evolve to meet the dynamic requirements of hotel management and staff.

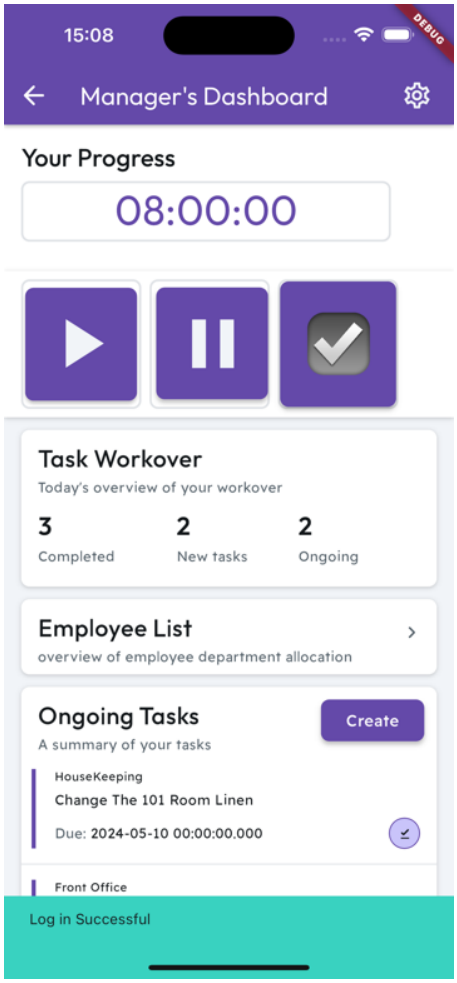


Fig.1. WORKOVER Dashboard

V. DELIVERABLES

The deliverables for the mobile application development project are designed to enhance the efficiency and effectiveness of hotel operations. The first deliverable involves creating a feature within the application that facilitates the seamless creation and management of shift handovers. This feature will allow employees to document essential information, such as tasks completed, pending tasks, and relevant notes, ensuring a smooth transition between shifts.

The second deliverable focuses on developing a functionality that enables managers to allocate tasks to specific departments within the hotel. This feature will ensure that tasks are assigned to the appropriate teams, streamlining workflow and improving overall efficiency.

The third deliverable involves creating a feature that allows employees to accept tasks assigned to them. This feature will provide a clear overview of the tasks assigned to each employee and their acceptance status, thereby facilitating better coordination and communication.

The fourth deliverable includes developing a shift timer feature that tracks the duration of each shift. Additionally, it will calculate the progress of tasks during the shift, providing real-time updates on task completion and overall shift progress.

The fifth deliverable focuses on creating a reporting feature that allows users to generate and access all task-related documents. This feature will ensure that important information, such as task lists, completion status, and notes, is easily accessible for reporting and analysis purposes.

The final deliverable involves developing a feature that allows administrators to allocate users to specific departments within the hotel. This will ensure that each user has access to relevant information and functionalities based on their departmental requirements.

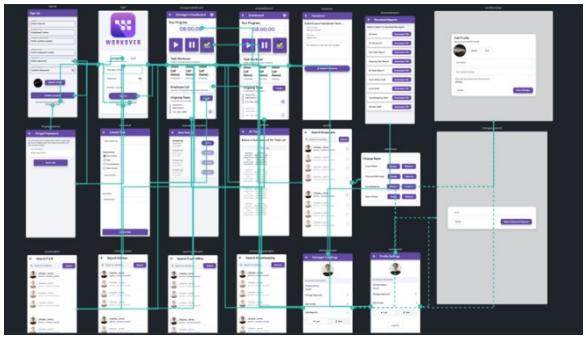


Fig.2. WORKOVER Application Flow

Cross-Platform Mobile UI: This deliverable focuses on creating a user interface that is compatible with multiple mobile platforms, such as iOS and Android. It will ensure that the mobile application provides a consistent and user-friendly experience across different devices, enhancing usability and accessibility.

VI. Discussion and conclusion

The development of the WORKOVER mobile application for hotel operations, as discussed, targets critical areas such as shift handovers, task allocation, and real-time reporting, which are essential for enhancing operational efficiency and communication within the hotel environment. The features designed for this application directly address the challenges identified in the preliminary phases of the research, such as inefficiencies in task management, miscommunication, and the lack of real-time updates.

The Shift Handover Creation feature is pivotal in ensuring seamless transitions between shifts by documenting completed tasks, pending tasks, and critical notes. This minimizes the risk of miscommunication and ensures continuity in operations, thereby fostering a cohesive and efficient workflow. The ability to streamline shift handovers directly impacts the quality of service provided to guests, as it reduces the likelihood of errors and delays.

Similarly, the Task Allocation to Departments feature is instrumental in optimizing workflow by ensuring that tasks are assigned to the appropriate teams. This structured allocation not only enhances efficiency but also helps in balancing the workload across departments, ensuring that resources are utilized effectively without overburdening any specific team. By streamlining this process, the application contributes to a more organized and productive working environment.

The Task Acceptance feature enhances accountability and communication by allowing employees to accept assigned tasks and providing real-time updates on task status. This transparency ensures that managers can track task progress effectively, facilitating better coordination and timely adjustments as needed. The clear delineation of responsibilities and the ability to monitor task completion status in real-time are crucial for maintaining operational efficiency.

The Shift Timer and Progress Calculation feature offers a robust mechanism for monitoring shift performance, providing managers with real-time data on task progress and shift duration. This feature is essential for identifying potential bottlenecks and ensuring that tasks are completed within the expected timeframe. By enabling managers to make data-driven decisions, this feature supports the continuous improvement of workflow processes.

The Reporting of All Task-Related Documents feature plays a vital role in maintaining comprehensive records of all activities, which is critical for both immediate operational needs and long-term strategic planning. By ensuring that all task-related information is easily accessible, the application supports transparency and accountability, making it easier to review and analyze operational data.

Finally, the Users and Department Allocations feature personalizes the user experience by providing employees with access to the tools and information relevant to their specific roles. This targeted approach improves efficiency by reducing information overload and ensures that staff members are focused on their designated tasks, thereby enhancing overall productivity.

The research project successfully demonstrates the potential impact of the mobile application on hotel operations. By addressing key challenges related to shift handovers, task allocation, and reporting, the application is poised to significantly enhance operational efficiency and employee productivity. The strategic alignment of these features with the long-term objectives of the hotel industry highlights the importance of a data-driven, transparent, and organized approach to hotel management.

The successful implementation of the application's features will not only improve day-to-day operations but also contribute to sustained improvements in how hotels manage their workforce. The reduction in manual errors, improved communication, and comprehensive reporting capabilities will support better decision-making and ultimately lead to higher guest satisfaction. In conclusion, this mobile application represents a significant step forward in modernizing hotel operations, making them more resilient, efficient, and responsive to the dynamic demands of the hospitality industry.

References

[1] M. Hort, M. Kechagia, F. Sarro, and M. Harman, “A Survey of Performance Optimization for Mobile Applications,” IEEE Transactions on Software Engineering, vol. 48, no. 8, pp. 2879–2904, 2022. [Online]. Available: https://doi.org/10.1109/TSE.2021.3071193. [Accessed: 28-Nov-2023].

[2] "I-centric communications: personalization, ambient awareness, and adaptability for future mobile services," IEEE Journals & Magazine. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/1336722. [Accessed: 30-Nov-2023].

[3] E. Kangas and T. Kinnunen, “Applying user-centered design to mobile application development,” Communications of the ACM, vol. 48, no. 7, p. 55, 2005. [Online]. Available:

https://doi.org/10.1145/1070838.1070866. [Accessed: 02-Dec-2023].

[4] "Shift handover," Fatfinger, SEE Forge creators of FAT FINGER. [Online]. Available: https://fatfinger.io/shift-handover/. [Accessed: 02-Dec-2023].

[5] "Multi-task allocation in mobile crowd sensing with ...," IEEE Xplore. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/8259371. [Accessed: 04-Dec-2023].

[6] "Designing an agile methodology for mobile software development: A hybrid method engineering approach," 2008. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/4632123. [Accessed: 03-Feb-2024].

[7] R. Payne, "Using Firebase with Flutter," in Apress eBooks, 2019, pp. 255–285. [Online]. Available: https://doi.org/10.1007/978-1-4842-5181-2_12. [Accessed: 12-Feb-2024].

[8] "Performance and stability Comparison of React and Flutter: Cross-platform Application Development," 2022. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/9996039. [Accessed: 25-Feb-2024].

[9] R. R. Prayoga et al., "Performance analysis of BLOC and Provider State Management Library on Flutter," 2021. [Online]. Available: https://www.ejournal.iocscience.org/index.php/mantik/article/view/1539. [Accessed: 28-Feb-2024].

[10] "Securing Flutter Applications: A Comprehensive study," 2023. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/10392001. [Accessed: 02-Mar-2024].

[11] T. Tran, "Flutter native performance and expressive UI/UX," 202 [Online]. Available: https://www.theseus.fi/handle/10024/336980. [Accessed: 10-Mar-2024].

[12] A. Kaur, "App review: Trello," Journal of Hospital Librarianship, vol. 18, no. 1, pp. 95–101, 2018. [Online]. Available: https://doi.org/10.1080/15323269.2018.1400840. [Accessed: 10-Mar-2024].

[13] J. Kim, D. J. Connolly, and S. C. Blum, "Mobile Technology: An Exploratory Study of Hotel Managers," International Journal of Hospitality & Tourism Administration, vol. 15, pp. 417–446, 2014. [Accessed: 29-Nov-2023].

[14] D. D. Patel, A. Khirsariya, and J. Shah, "Automated system for hotel room service," Semantic Scholar. [Online]. Available: https://www.semanticscholar.org/paper/Automated-system-for-hotel-room-service-Patel-Khirsariya/bab2b7c0b86800483ddebb0739dfcdfbda4aafaa. [Accessed: 29-Nov-2023].

[15] S. Kadry, A. Bagdasaryan, and M. Kadhum, "Simulation and analysis of staff scheduling in hospitality management," in 2017 7th International Conference on Modeling, Simulation, and Applied Optimization (ICMSAO), 2017. [Online]. Available: https://www.semanticscholar.org/paper/Simulation-and-analysis-of-staff-scheduling-in-Kadry-Bagdasaryan/c9af3cf885f69d9090425d439c01307337f800ca. [Accessed: 30-Nov-2023].

[16] M. Walerych and W. Zabierowski, "Application supporting hotel management as an example of web technologies usage," in Proceedings of International Conference on Modern Problem of Radio Engineering, Telecommunications and Computer Science, pp. 327-328, 2012. [Accessed: 30-Nov-2023].

[17] B. Bullock, C. Learmonth, H. Davis, and A. Al Mahmud, "Mobile phone sleep self-management applications for early start shift workers: A scoping review of the literature," Frontiers in Public Health, vol. 10, 2022. [Accessed: 30-Nov-2023]

[18] H. K. Flora and S. V. Chande, "A review and analysis on mobile application development processes using agile methodologies," International Journal of Research in Computer Science, vol. 3, no. 4, p. 9, 2013.

NStudy Study Room Management System for NSBM Green University

Onaliy Jayawardana
Faculty of Computing
NSBM Green University
Homagama

Kalindu Perera
Faculty of Computing
NSBM Green University
Homagama

Geeth Induwara
Faculty of Computing
NSBM Green University
Homagama

Avish Rodrigo
Faculty of Computing
NSBM Green University
Homagama

Jayamuni Rashminda
Faculty of Computing
NSBM Green University
Homagama

***Abstract*—This research, discusses about the issues faced by students and administration with the manual study room booking process at NSBM Green University and this research proposes an innovative IoT system to enhance the management of those study rooms. The system aims to simplify the booking process for students through an intuitive online platform to create a smart infrastructure. The primary goal is to improve student experiences by providing safe, automated access to their booked study rooms during designated times, thus ensuring energy efficiency and promoting sustainability. Additionally, the system monitors noise levels within study rooms to maintain a conducive study environment, alerting students to disturbances. Real-time monitoring and data collection enable the system to provide insights into environmental factors and room usage trends, facilitating long-term resource optimization and informed decision-making.**

Keywords—IoT, Study rooms, Sensors, Micro controllers, Real time data

I. INTRODUCTION

The management of study rooms within educational institutions, including those at National School of Business Management Green University, is hindered by several challenges stemming from manual processes and outdated systems. Currently, the booking process and overall management of study rooms at National School of Business Management Green University rely solely on manual procedures, including paper-based booking forms and physical key distribution.

This manual approach results in inefficiencies in room utilization, as students often encounter difficulties in securing study room bookings. Moreover, the lack of a centralized and automated system for access control leaves study rooms vulnerable to unauthorized usage and security breaches.

In addition to security concerns, the manual management of study rooms also presents challenges in maintaining a conducive learning environment. Without real-time monitoring capabilities, administrators struggle to address disturbances such as noise promptly, impacting the quality of the study environment and student experiences.

Furthermore, the absence of data insights hampers administrators' ability to make informed decisions regarding room usage optimization and energy consumption. Current systems lack the granularity and responsiveness needed to accurately track occupancy levels and monitor environmental factors in real-time.

Given these challenges, there is an urgent need for a comprehensive IoT-based study room management system at NSBM Green University. Such a system should streamline the booking process, automate access control using Radio Frequency Identification (RFID) technology, and incorporate advanced sensor technologies for real-time monitoring of occupancy and noise levels.

Moreover, the system should provide administrators with a user-friendly interface to visualize and analyze data, enabling informed decision-making and enhancing the overall learning experience for students. By addressing the limitations of manual processes and outdated systems, the proposed solution aims to create a more efficient, secure, and conducive learning environment at National School of Business Management Green University.

II. BACKGROUND STUDY

Various systems and technologies have been developed to address challenges related to study room management and environmental monitoring in educational institutions. These systems offer insights into existing approaches and technologies utilized in similar contexts.

A. Radio Frequency Identification (RFID) – Based Control Systems

Radio Frequency Identification (RFID) (Radio Frequency Identification) technology has been widely adopted for access control in various settings, including educational institutions. These systems utilize Radio Frequency Identification (RFID) cards or tags to grant access to authorized individuals. For example, libraries often use Radio Frequency Identification (RFID)-based systems to manage access to study rooms or restricted sections. These systems typically involve Radio Frequency Identification (RFID) readers installed at entry points, which communicate with Radio Frequency Identification (RFID) cards carried by users to verify identity and grant access. [1]

B. Occupancy Sensing Systems

Occupancy sensing systems employ various technologies, such as infrared sensors, ultrasonic sensors, and motion detectors, to detect the presence of individuals in a given space. In educational institutions, occupancy sensing systems are often used to monitor classroom utilization, study room occupancy, and library seating availability. These systems enable administrators to optimize room allocation, enhance space utilization, and improve energy efficiency by controlling lighting and HVAC systems based on occupancy status. [1]

C. Integrated IoT-Based Study Room Management Systems

Recent advancements in IoT (Internet of Things) technology have paved the way for integrated study room management systems that combine Radio Frequency Identification (RFID)-based access control, occupancy sensing, noise detection, and data analytics capabilities. These systems leverage sensors, wireless communication technologies, and cloud-based platforms to provide real-time monitoring, intelligent automation, and data-driven insights. By integrating multiple functionalities into a unified platform, IoT-based study room management systems offer enhanced flexibility, scalability, and efficiency compared to standalone solutions.

III . LITERATURE REVIEW

This literature review aims to explore existing studies and projects addressing the implementation of smart study rooms, focusing on Radio Frequency Identification (RFID) integration, occupancy sensing noise detection embedded system.

D. Radio Frequency Identification (RFID) Integration and Access Control

Numerous studies have explored the use of Radio Frequency Identification (RFID) technology for access control. Fatah Chetouane implemented an Radio Frequency Identification (RFID)-based system for access to a library, utilizing student ID cards. The system demonstrated efficiency in authentication; however, limitations were identified in terms of user granularity and real-time access management. [2]

E. Occupancy Sensing in Smart Buildings:

In the realm of occupancy sensing, Yudith Cardinale developed a smart building management system using infrared sensors. While their approach effectively detected occupancy, issues related to sensor accuracy and false positives were noted. [3]

F. Noise Detection and Alert Systems:

Research by Shaikh Shahrukh India Mulani Rameez, A.C.Patil and Khan Abdul (2021) focused on noise detection. The study provided valuable insights into noise anomaly detection; however, the responsiveness of their alert system raised concerns. [4]

G. Comparative References and Performance Benchmarks

To enhance the theoretical rigor of this study, incorporating comparative references and performance benchmarks is essential. Analyzing existing solutions and their respective performance metrics, such as response times, accuracy rates, and user satisfaction levels, will provide a comprehensive understanding of the effectiveness of different technologies.

For instance, a comparative analysis of various occupancy sensors employed in educational settings could elucidate the advantages and drawbacks of each approach. Similarly, referencing case studies that highlight the performance of RFID access control systems in other institutions will offer insights into potential scalability and adaptability of the proposed NStudy system.

Additionally, establishing benchmarks for energy efficiency improvements and noise reduction could substantiate the anticipated benefits of the NStudy system, validating its effectiveness through empirical data.

III. PROPOSED SOLUTION

The solution to the identified problem lies in the implementation of an innovative IoT system called NStudy study room management system. This system is designed to revolutionize the management of study rooms at National School of Business Management Green University by introducing automation, real-time monitoring, and data-driven decision-making processes.

A. Intuitive Online Platform

NStudy study room management system will offer an intuitive online platform that simplifies the process of booking study rooms for university students. Through this platform, students can easily check room availability, select suitable time slots, and make bookings with minimal effort. All the booking details are stored in the Firebase Realtime Database. Also, conflicting bookings are avoided and also, they are recorded in the Firebase Realtime Database for decision making purposes. The user-friendly interface enhances accessibility and encourages more students to utilize the available study spaces efficiently.

B. Smart Infrastructure:

The core of the NStudy study room management system will be its smart infrastructure, which integrates various IoT components such as ESP32 Microcontroller, human presence sensor, sound sensor, light sensor, Radio Frequency Identification (RFID) scanner, solenoid lock, LCD Panel, buzzer, and cloud connectivity. These components work together seamlessly to create a connected environment that enhances the management of study rooms in several ways.

C. Automated Access

By leveraging solenoid lock and Radio Frequency Identification (RFID) scanner, NStudy study room management system will provide students with safe and automated access to the study rooms they have booked during designated time intervals. This automation eliminates the need for manual intervention, ensuring a smooth and hassle-free experience for students.

D. Energy Efficiency

NStudy study room management system will promote energy efficiency by only activating electricity in study rooms when they are occupied. Also, the if a human presence is not detected, the electricity will turn off. By dynamically adjusting room conditions based on occupancy status, the system will minimize energy wastage and contribute to sustainability efforts within the university.

E. Noise Monitoring

The inclusion of sound sensors will allow NStudy study room management system to monitor noise levels within study rooms in real-time. In the event of excessive noise or disturbances, the system will promptly alert students by buzzing the buzzer inside the study room, enabling timely intervention to maintain a conducive study environment.

F. Data Driven Insights and Predictions

NStudy study room management system will go beyond basic room management by leveraging cloud connectivity to collect and analyze a wealth of data on environmental factors and room usage patterns. This data-driven approach will enable administrators to gain valuable insights into resource utilization and occupancy trends, empowering them to make informed decisions for optimizing room allocation and enhancing overall efficiency. Predictions will be provided for peak times and days, which will help administration to allocate more study rooms at those peak times and days.

Accordingly, the core challenges we seek to overcome with NStudy study room management system are as follows:

A. Manual Reservation Process: Transitioning from a paper-based logbook system to a digitized, online reservation platform.
B. Security and Access Control: Ensuring the integrity of student IDs, preventing unauthorized access, and implementing secure access control mechanisms.
C. Administrative Efficiency: Minimizing administrative overhead, reducing errors, and optimizing the utilization of study rooms.
D. User Experience: Empowering students with a user-friendly and convenient reservation system accessible from any device.

In summary, the proposed NStudy IoT study room management system will offer a comprehensive solution to the challenges faced in study room management at NSBM Green University. By combining an intuitive online platform, smart infrastructure, and data-driven insights, NStudy study room management system will not only improve student experiences but also will promote sustainability and enable administrators to make informed decisions for a more adaptable and sustainable learning environment. By adopting to NStudy study room management system, National School of Business Management can aim to revolutionize its study room management practices, fostering a secure, efficient, and user-centric environment conducive to optimal learning experiences.

IV. THEORETICAL FOUNDATION

The theoretical foundation of the NStudy IoT study room management system is grounded in concepts of smart environments, IoT architecture and sustainability in education.

A. Smart Environments

This concept emphasizes the integration of technology to create responsive and adaptive environments that improve user experience. By employing IoT technologies, NStudy study room management system transforms traditional study room management into a smart, user-centric system that caters to the needs of students and administrators alike.

B. IoT Architecture

The system follows a layered architecture model, encompassing sensing, networking, data processing,

and application layers. Each layer plays a critical role in ensuring the seamless operation of the system. Sensing devices gather real-time data, networking components transmit this data to the cloud, where it is processed and analyzed, and finally, applications provide users with meaningful insights and functionalities.

C. Sustainability in Education

The NStudy IoT study room management system aims to support sustainability efforts within educational institutions. By optimizing resource usage, enhancing energy efficiency, and providing a conducive learning environment, it contributes to broader sustainability goals that many educational institutions are striving to achieve.

V. TECHNOLOGIES

A. Radio Frequency Identification (RFID) Technology

Radio Frequency Identification (RFID) uses electromagnetic fields to identify and track tags attached to objects. In our study room system, RFID readers are installed at the room entrances to detect RFID-enabled cards carried by users. Each RFID card contains a unique identifier that is read by the RFID reader, allowing for secure access control. When a user swipes their card, the system authenticates the card's unique ID against the stored data in **Firebase**, ensuring that only authorized users can access the room. This integration facilitates seamless and contactless room access for students, enhancing both convenience and security.

B. Sound Sensor

The sound sensor continuously monitors the noise levels within the study room by detecting sound waves and converting them into electrical signals. This component helps maintain a conducive study environment. If the detected sound level exceeds a predefined threshold (e.g., due to loud conversations or disturbances), the system triggers a **buzzer** and displays a warning on the **LCD screen**. This immediate feedback allows students to take corrective actions and ensures that the study room remains quiet and focused.

C. Buzzer

The buzzer serves as an auditory alert system. When the sound sensor detects noise levels beyond the acceptable range, the buzzer is activated to notify users. It provides immediate and clear feedback, alerting students when noise levels are too high, which is crucial for maintaining an optimal study environment. The buzzer also works in tandem with real-time notifications displayed on the **web application** to enhance situational awareness.

D. Smart Human Presence Sensor

Smart human presence sensors detect the presence of people within a specific are. In this proposed study room management system, these sensors will detect human presence within the room, allowing the system to automatically activate or deactivate lights accordingly. This helps optimize energy usage and enhances study room functionality.

E. Light Sensor

Light sensors, also known as photodetectors, measure ambient light levels in their surroundings. In our system, light sensors will monitor the amount of natural and artificial light present in the study room. This information can be used to know whether light is turned on or off during unwanted time periods, which helps to optimize energy consumption.

F. LCD Display

An LCD (Liquid Crystal Display) is a flat panel display technology that uses liquid crystals to produce text. In our study room system, an LCD display will provide real-time feedback and information to users. It can display various parameters, such as room occupancy status, system alerts, and instructions for operating the system, enhancing user interaction and convenience.

G. Solenoid Lock

A solenoid lock is an electromechanical locking device that operates by using an electrical current to move a locking bolt or latch. In our study room system, a solenoid lock will control access to the room by electronically locking and unlocking the door in response to valid Radio Frequency Identification (RFID) authentication. This provides secure access control and prevents unauthorized entry into the study room.

H. Relay Module

A relay module is an electromechanical switch that is used to control the flow of electricity to various electrical devices and appliances. In our proposed study room management system, a relay module will interface with the microcontroller to control the operation of the solenoid. This allows for the solenoid lock to be activated and deactivated based on the signals received from the microcontroller, ensuring smooth and efficient operation of the access study room.

I. 12V Power Supply

The 12V power supply provides the necessary electrical power to operate the components of the smart study room management system. It ensures reliable and stable operation of the system, providing the necessary power for solenoid lock through a relay module to function the study room system efficiently.

J. ESP32 Micro-controller

The ESP32 is a powerful microcontroller with built-in Wi-Fi and Bluetooth capabilities. It's commonly used in IoT applications; we propose to use the ESP32 as the microcontroller in this proposed study room management systems for its versatility, low power consumption, and ease of integration with various sensors and communication protocols.

K. Firebase

Firebase is a mobile and web application development platform that provides various services, including real-time database, authentication, hosting, and cloud functions. In this proposed study room management system, Firebase can be used for data storage, user authentication, and real-time communication between the study room system and remote devices, enabling seamless access control, data management, and monitoring capabilities.

L. Web Application

A web application built using React will create a user-friendly interface for managing study room bookings, displaying real-time data from the system, and allowing users to interact with the system remotely. And admins can view data related to the proposed study room management system, e.g., peak usage times of the study room.

M. Arduino IDE

Arduino IDE is open-source software used to program Arduino microcontrollers. It provides a simple and intuitive interface for writing, compiling, and uploading code to Arduino boards. We propose to use the Arduino IDE to integrate hardware components into this proposed study room management system.

Using these technologies, we were able to create a seamless experience for users to easily book study rooms and for admins to efficiently monitor usage patterns. This integration of React and Arduino IDE allowed us to streamline the process of managing study room bookings and improve overall system functionality. Additionally, to create an efficient study room system that enhances security, comfort, and productivity for users.

VI. Methodology

The methodology for the above proposed solution of developing the NStudy IoT study room management system involves several key phases, including system design, component selection, solution implementation and solution evaluation.

A. System Design

Under the system design phase, conducting a requirement analysis through surveys, interviews and questionaries with students and university administration to understand their needs, requirements and expectations regarding the study room management system will be done.

In addition to the above, developing system architecture diagrams to outline the interaction between various IoT components, software interfaces and the cloud infrastructure will be done.

B. Component Selection

Under this phase, choosing the appropriate IoT devices which are mentioned in the above technologies section based on their functionality, reliability and cost effectiveness will be conducted.

C. Implementation

Under this phase, the web application platform will be developed using React as explained in the Web Application sub section of technologies section. The front end of the web application will be developed using React while integrating Firebase for real time database functionalities.

D. Evaluation

Under this phase a pilot testing will be conducted for the Nstudy IoT study room management system, with a selected group of students to gather feedback on usability, functionality and overall experience.

Also, analyze performance metrics such as booking success rates, user engagement levels, energy consumption data and noise disturbance reports to assess the effectiveness of the system will be conducted in this phase.

VII. Outcome and In-Depth Analysis of Results

The development of the NStudy IoT study room management system is expected to bring about several positive outcomes, revolutionizing study room management at NSBM Green University and enhancing the overall student experience. The following are the anticipated outcomes based on the design and functionality of the NStudy study room management system:

A. Streamlined Booking Process

The implementation of the intuitive online platform will simplify the booking process for study rooms, making it more accessible and user-friendly for students. Also, it will avoid booking collisions and provides a better visualization of study room availability for students.

Expected metrics: We expect more than 90% decrease in booking conflicts within the first month of implementation of the system.

B. Enhanced Efficiency and Accessibility

The automation of access to study rooms through solenoid locks and Radio Frequency Identification (RFID) scanners is anticipated to improve efficiency and accessibility, reducing the need for manual interventions and enhancing convenience for students.

This ensures only authorized access to the study rooms.

C. Improved Resource Utilization

NStudy study room management system's smart infrastructure, including its ability to activate electricity only when rooms are occupied, leads to more efficient utilization of study room resources, contributing to cost savings and sustainability. Similar systems have reported energy savings ranging from 15% to 30%.

Expected metrics: Energy consumption data analyzed pre- and post-implementation to quantify savings.

D. Maintaining a Conducive Study Environment

The real-time monitoring of noise levels within study rooms helps to maintain a conducive study environment by promptly addressing noise-related issues and disturbances. The buzzer being buzzed at high noise levels alerts students to be quiet.

Expected Metrics: Surveys assessing student satisfaction regarding noise levels and overall study conditions, aiming for improved ratings compared to previous experiences.

E. Data Driven Insights and Predictions

The collection and analysis of data on room usage patterns are expected to provide administrators with valuable insights for optimizing room allocation and enhancing operational efficiency.

Expected Metrics: Administrative reports detailing usage patterns, peak times, and suggestions for resource reallocation.

F. Cost Effective and Sustainable Solution

NStudy study room management system is designed to offer a cost-effective solution to study room management challenges, utilizing existing infrastructure and IoT technologies to improve efficiency and student satisfaction while promoting sustainability.

VIII. Critical Analysis of Proposed Outcomes

A. Feasibility Assessment

An essential component of this proposal is a critical analysis of the feasibility of implementing the NStudy IoT study room management system. Factors such as the budget required for implementation, technical requirements, and potential challenges need to be considered. For example, the integration of various IoT devices may necessitate specialized skills for installation and maintenance.

B. Impact Analysis

A thorough impact analysis should address both the anticipated benefits and potential drawbacks of implementing the NStudy IoT study room management system. The impact on different stakeholders, including students and administrators, must be evaluated. For instance, the increased efficiency in room booking may enhance the user experience for students, while the data-driven insights could empower administrators to make more informed decisions regarding resource allocation.

C. Risk Factors

Identifying potential risks associated with implementation is crucial. These may include technology failures, user adoption challenges, and security concerns related to student data. Proposing mitigation strategies, such as training for users and ensuring robust cybersecurity measures, will be vital for successful implementation.

References

[1] X. Jia, Q. Feng, T. Fan and Q. Lei, " RFID Technology and Its Applicationsin Internet of Things (IOT)," 2012.

[2] F. Chetouane, "An Overview on RFID Technology Instruction and Application," in International Federation of Automatic Control, 2015.

[3] Y. Cardinale, "Occupant Activity Detection in Smart Buildings: A Review," 2020.

[4] S. Shahrukh, M. Rameez and K. Abdul, "Noise Detection System," International Journal of Engineering Applied Sciences and Technology, 2021.

Optimizing the Coconut Supply Chain: Implementing a Blockchain and IoT-Integrated Management System for Sustainable Agritech in Sri Lanka

Hiranya Thrimawithana
Department of Software Engineering & Information Systems
NSBM Green University
Homagama, Sri Lanka
htthrimawithana@students.nsbm.ac.lk

Ms. Lakni Peiris
Department of Software Engineering & Information Systems
NSBM Green University
Homagama, Sri Lanka
lakni.p@nsbm.ac.lk

Abstract— **This research investigates the improvement of supply chain management (SCM) in Sri Lanka's coconut industry by incorporating information systems and blockchain technology. It offers CocoGuard, a system that leverages smart contracts and IoT sensors to increase transparency, efficiency, and sustainability. Findings from case studies and surveys related to the Coconut Triangle show that CocoGuard improves real-time data monitoring and automates transactions, resulting in more effective resource management and waste reduction. Its scalable and adaptable design suggests broad applicability across various agricultural sectors, fostering sustainable practices and transforming traditional SCM methods.**

Keywords—Coconut industry, Blockchain, IoT Sensors, Supply Chain Management, Sustainability, RFID

I. INTRODUCTION

In the past, agriculture was the cornerstone of global economies, serving as the primary sector on which societies depended for survival. It provided food and resources, supporting the livelihoods of countless individuals and communities. Agriculture's importance is deeply rooted in human civilization, evolving from simple farming methods to a complex, interconnected system that drives economic growth and sustains populations. Across various countries, agriculture remains a vital sector, contributing significantly to national economies. For instance, in Indonesia, agriculture is a key economic driver and a major source of livelihood for the population. Similarly, in India, agriculture contributes 18% to the Gross Domestic Product (GDP) and employs 50% of the workforce, making it a crucial pillar of the economy. India leads the world in producing pulses, rice, wheat, spices, and spice products.[1]

The agricultural sector is diverse, encompassing various subsectors. Among them, the crop sector plays a pivotal role, with the coconut industry standing out due to its versatility and economic significance. Coconut cultivation spans over 90 countries, predominantly in Asia, the Pacific Islands, and South America. The global planted area for coconuts covers approximately 12 million hectares, yielding an estimated 70 billion nuts annually.[2] Traditional coconut products include coconut oil, copra meal, desiccated coconut, and coir, while newer products like coconut water, virgin coconut oil, and coconut milk are rapidly gaining popularity.[2]

India, Indonesia, and the Philippines are the leading coconut producers, collectively accounting for over 75% of global production. Sri Lanka, Brazil, Papua New Guinea,

Vietnam, Mexico, Thailand, and Malaysia are also significant contributors. In 2014, global coconut production reached 69.8 billion nuts, maintaining a steady output of between 65 and 70 billion nuts annually from 2010 to 2014.[2] Currently (2024), worldwide coconut production is about 62.4 million tons, highlighting the important role of the crop in global agriculture.

In Sri Lanka, the coconut industry is a cornerstone of both the agricultural sector and the broader economy. Coconuts are integral to Sri Lankan cuisine, providing essential ingredients like coconut milk and oil. After rice, coconut is the second most demanded food in Sri Lanka, accounting for 10% of the total agricultural produce.[3] Beyond its culinary uses, coconut is also vital in cosmetics, pharmaceuticals, and handicrafts, underscoring its economic importance. Coconut cultivation occupies approximately 20% of Sri Lanka's arable land, totaling nearly 440,457 hectares.[3] The sector contributes 0.7% to Sri Lanka's GDP and represents nearly 5% of the country's export earnings.[3]

The problem in Sri Lanka's coconut industry is its fragmented supply chain, lacking transparency, coordination, and efficiency. This results in poor communication, delays, and increased costs across stakeholders. The motivation for the study stems from the industry's economic significance and the need to address these inefficiencies to boost sustainability and global competitiveness. The research objectives are identifying key supply chain issues, assessing their impact on efficiency and sustainability, and developing a customized SCM system to improve the industry's performance.

However, the Sri Lankan coconut industry faces numerous supply chain challenges which are

inadequate infrastructure, logistics issues, and coordination problems among stakeholders, all hindering coconut products' efficient production and distribution. Addressing these challenges through a streamlined Supply Chain Management (SCM) system is crucial for enhancing the industry's competitiveness and sustainability. A robust SCM system can improve transparency, traceability, and coordination, ultimately ensuring that high-quality coconut products meet market demands sustainably.

II. LITERATURE REVIEW

Supply Chain Management (SCM) is a critical process that involves coordinating and overseeing a network of interconnected businesses to ensure the efficient flow of goods and services from suppliers to end customers. SCM encompasses all activities related to sourcing, procurement, production, and logistics, integrating supply and demand management both within and across companies. The importance of SCM is evident across various industries, including manufacturing, healthcare, apparel, and logistics, each of which adapts SCM to meet its specific needs.

In manufacturing, SCM plays a vital role in optimizing inventory management, production scheduling, and supplier relationships. Practices like Just-in-Time (JIT) and Lean Manufacturing are widely adopted to minimize waste and enhance efficiency. [4] In the healthcare industry, SCM is crucial [5] for managing medical supplies, cold chain logistics, and regulatory compliance, ensuring the safety and efficacy of medical products. The apparel industry relies on SCM [6] for fast fashion inventory management and ethical sourcing, enabling quick adaptation to fashion trends while maintaining sustainability. Meanwhile, in the logistics industry, SCM is essential [7] for transportation management, warehouse management, and third-party logistics (3PL) services [8], ensuring the smooth movement and storage of goods.[9]

The agricultural sector, particularly, has undergone significant transformations in SCM practices [10]. Traditionally, agricultural supply chains were localized and linear [11], with goods moving directly from farmers to nearby markets or consumers. Technology played a minimal role, and communication and coordination relied heavily on direct relationships and local markets. Seasonal variability greatly influenced production and supply, with limited infrastructure for storage and processing. However, the advent of technology has revolutionized agricultural SCM [12], expanding it from traditional to technology-based practices. This shift has brought about a heightened focus on the environmental, social, and economic impacts of food production and consumption. As a result, there is increasing pressure on consumer organizations, social and environmental advocacy groups, agricultural organizations, and policymakers to create sustainable supply chains.[9]

Despite the advancements in SCM, the agricultural industry still faces significant challenges [13], particularly in developing countries. In the context of the agricultural food supply chain, major challenges include the lack of industrialization, poor management practices, inaccurate information, and inefficient supply chains [14]. These issues are further exacerbated by the increasing global population and the subsequent rise in demand for agricultural products [15]. To address these challenges, advanced technological trends such as blockchain, the Internet of Things (IoT), and Artificial Intelligence (AI) have been integrated into SCM systems.[9]

Blockchain and IoT technologies have had a profound impact on various industries [16], including agriculture, finance, healthcare, and logistics. In SCM, these technologies enhance transparency, efficiency, and security. IoT devices gather and transmit real-time data on the condition and location of goods, while blockchain stores this data in an immutable ledger, ensuring accurate, transparent, and secure information. The integration of blockchain and IoT in SCM offers significant benefits, including enhanced traceability, accountability, and efficiency.[17]

In the agricultural sector, blockchain technology addresses several challenges faced by SCM stakeholders [18], particularly in ensuring the safety and quality of agricultural products. With the growing complexity of agricultural supply chains, there is an increasing demand for transparent and reliable methods to trace the origin and authenticity of food products. Blockchain technology provides a secure and immutable record of transactions, ensuring that products meet the required standards. Moreover, the integration of IoT devices allows for real-time data collection and monitoring, further enhancing the efficiency and effectiveness of SCM systems.[17]

The combination of blockchain and IoT technologies also offers substantial benefits to stakeholders throughout the supply chain. These technologies improve productivity and efficiency while reducing the risk of fraud and counterfeiting. By implementing blockchain technology in agriculture, stakeholders can improve safety, reduce environmental impact, and increase profits. IoT technology, in particular, reduces the workload for farmers and other stakeholders by automating data collection and transmission, thereby enhancing data accuracy and transparency throughout the supply chain.[19]

Despite these advancements, the Sri Lankan coconut industry still lacks a comprehensive SCM system that utilizes these technologies [20]. The traditional supply chain processes in the industry are plagued by inefficiencies, lack of transparency, and inadequate management practices. Dependence on outdated methods and the absence of industrialization have impeded industry’s ability to meet growing market demands sustainably. Without the integration of blockchain and IoT technologies, the Sri Lankan coconut industry struggles with challenges related to traceability, quality control, and logistics efficiency. Addressing these issues through the adoption of

advanced SCM technologies is crucial for enhancing the productivity, sustainability, and profitability of the coconut industry in Sri Lanka.[21]

III. Research Methodology

This research methodology uses a systematic framework that combines qualitative and quantitative research techniques through a mixed methods approach. This approach not only quantifies supply chain efficiency but also captures stakeholder challenges through interviews.

A. Research Design

This study employs a mixed-methods approach, combining both qualitative and quantitative research designs to examine in-depth the challenges and opportunities within Sri Lanka's coconut supply chain. The qualitative component includes in-depth case studies and interviews with key industry stakeholders – farmers, manufacturers, distributors, retailers, and consumers. This helps to identify current inefficiencies and gaps in the supply chain and gathers information on stakeholder views on potential technological solutions.

On the other hand, the quantitative component involves the collection and analysis of numerical data related to production, distribution, and sales in the coconut industry. This data is crucial to assess the extent of supply chain issues and evaluate the potential impact of proposed system improvements. Additionally, surveys are used to gather generalized feedback from stakeholders on the effectiveness of existing practices and their willingness to adopt new technologies such as blockchain and IoT.

B. Data Collection

Data collection is carried out in two stages: primary and secondary. Primary data is obtained through field research methods such as surveys, interviews, and focus group discussions with stakeholders in the Sri Lankan coconut industry. The surveys are designed to assess the current state of the supply chain, focusing on parameters such as production efficiency, traceability, transparency and stakeholder satisfaction. Respondents include farmers, manufacturers, distributors, retailers and consumers, providing a comprehensive perspective of the supply chain. Also, the survey is self-administered structured questionnaire. Survey is conducted by defining 7 sections with total number of 23 questions.

Interviews and focus groups with selected stakeholders offer deeper insights into the challenges they face and their views on potential technological solutions. These interviews are semi-structured, allowing for a combination of guided and open-ended questions, yielding detailed qualitative data. This information is essential to understanding the complexities of the supply chain and the specific needs of different stakeholder groups.

Secondary data is obtained from existing literature, industry reports, government publications and relevant databases. This includes information on the current state of Sri Lanka's coconut industry, global trends in agricultural supply chain management, and the implementation of blockchain and IoT in similar sectors. Secondary data provides a contextual basis for research and enhances the analysis of primary data by offering benchmarks and comparative insights.

C. Analysis Methods

The population for this study consists of coconut stakeholders in Sri Lanka, but due to the challenge of accessing all, a sample was selected using stratified sampling and snowball sampling techniques. The sample area was the Sri Lanka Coconut Triangle covering the districts of Puttalam, Kurunegala, and Gampaha. A mixed methods approach incorporating both quantitative and qualitative research methods was adopted to gain a deeper understanding of the research problem. For data analysis, thematic analysis was used for qualitative data, and descriptive analysis was applied for quantitative data, ensuring a comprehensive view of the findings.

IV. Results and Discussion

The study on "Optimizing the Coconut Supply Chain: Implementing a Blockchain and IoT-Integrated Management System for Sustainable Agritech in Sri Lanka" has provided valuable insights into the current supply chain situation and opportunities for improvement through modern technological solutions. These findings are drawn from qualitative data collected through interviews and focus group discussions with stakeholders, along with quantitative data obtained from surveys and system testing.

A. System Architecture

The CocoGuard SCM system features a modular design with four distinct layers: the SCM layer, the Blockchain layer, the IoT layer, and the Application layer. The SCM layer oversees supply chain operations, while the Blockchain layer uses smart contracts to automate and authenticate transactions, thereby promoting transparency and minimizing conflicts. The IoT layer integrates various sensors, such as RFID, GPS, temperature, soil moisture, pH level, humidity, and weight sensors, to collect real-time data on crucial supply chain metrics.

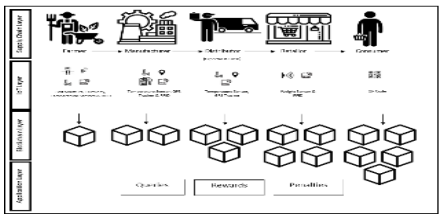


Fig 1: System Architecture

1) Blockchain Technology for Better Traceability

This research highlights the significant potential of blockchain technology to overcome traceability challenges identified by stakeholders. Smart contracts within blockchain technology play a crucial role in automating and enforcing agreements and conditions between stakeholders in a transparent and tamper-proof manner, greatly reducing errors and fraud within the supply chain. For example, a smart contract could automatically trigger payment to farmers once specific conditions such as product quality and delivery times are met. This not only improves efficiency but also builds trust between stakeholders by ensuring that all parties adhere to the agreed terms. This is especially important in the coconut industry, where traceability is vital to maintaining product quality and meeting the growing demand for sustainably sourced products.

2) IoT for Real-Time Data Collection

For real-time data collection, a variety of sensors are used to gather essential information throughout the supply chain. RFID tags enable detailed sensor data, while GPS devices provide precise location data. Temperature sensors assess environmental conditions, and soil moisture sensors monitor soil hydration. pH level sensors measure the soil pH level, humidity sensors assess atmospheric moisture, and weight sensors capture product mass. Together, these sensors provide a comprehensive data set that is critical to improving supply chain efficiency and ensuring product quality.

The following table explains the IoT sensors that are used in each stakeholder's stages.

TABLE I. TABLE 1: IOT SENSORS THAT HAVE BEEN USED

Level	**Variable**	**Rationale**
Farmer	Temperature Sensor	Monitors environmental temperature to ensure optimal growth conditions.
	Soil Moisture Sensor	Measures soil moisture to ensure proper irrigation.
	Humidity Sensor	Tracks air humidity to prevent diseases and manage plant health.
	ph Sensor	Measures pH levels of soil.
	RFID	Tracks and manages coconuts from the farm level.
Manufacturer	ph Sensor	Measures pH levels of processing environment and final products.
	Temperature Sensor	Monitors temperature during processing to meet safety standards.
	RFID	Tracks processed products and manages inventory.
Distributor	GPS	Tracks the location of shipments in real time.
	Temperature Sensor	Monitors temperature during transportation to ensure quality.
Retailer	Weight Sensor	Measures the weight of products for accurate inventory and sales tracking.
	RFID	Tracks products in the retail environment and manages stock.

The Application layer provides a user-friendly interface for stakeholders to interact with the system, view the data, and make informed decisions. These layers combined form a cohesive platform that improves efficiency, transparency, and sustainability within the Sri Lankan coconut industry.

The following diagrams display how the blockchain smart contract works. It displays the pre-conditions that each should have.

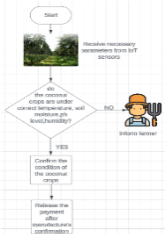


Fig. 2: Flow between Farmer & Manufacturer

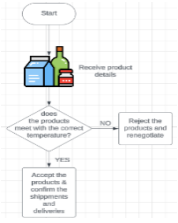


Fig 3: Flow Between Manufacturer and Distributor

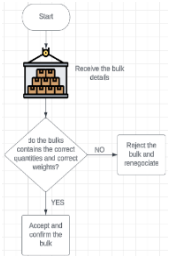


Fig 4: Flow Between Distributor & Retailer

B. System Design & Development

Building on insights gained from the data collection phase, research is moving forward with the design and development of a sustainable supply chain management (SCM) system tailored to Sri Lanka's coconut industry. The proposed system, called "CocoGuard," integrates blockchain technology to enhance transparency and traceability, and IoT for real-time data collection and monitoring. The current system is developed to integrate the functions of stakeholders (farmers and manufacturers) where each has multiple benefits from the proposed system and the smooth process has flowed throughout the system. Once the farmer Registered to the system he can use those credentials to log in and the farmer dashboard displays the respective year with the sensor details as well as the coconut harvested amount throughout the year. In the Manufacturer's dashboard who is the Admin of the system, can see the details of each production and the analysis of them as well as the current and expected revenue.

Fig. 5. Registration Page

Fig.7. Manufacturer Dashboard

C. Data Analysis

Data collected from surveys, interviews and system testing are examined using a combination of qualitative and quantitative approaches. Qualitative data from interviews and focus group discussions are subjected to thematic analysis to uncover recurring patterns, themes and perspectives related to challenges and opportunities within the coconut supply chain. This analysis provides a comprehensive understanding of stakeholder needs and preferences, which guides the design and implementation of the CocoGuard system.

On the other hand, quantitative data from surveys and system testing are analyzed using statistical methods to assess the effectiveness of current supply chain practices and the potential benefits of the CocoGuard system. Key performance indicators (KPIs) such as efficiency, transparency, traceability and stakeholder satisfaction are employed to measure the performance of the system.

Qualitative data collected from interviews and focus group discussions revealed important insights into the unique challenges faced by different stakeholders within the coconut supply chain. Farmers emphasized issues such as the absence of real-time data on crop conditions, which often leads to suboptimal harvest times and lower product quality. Manufacturers and distributors highlighted challenges in tracing the origin and quality of coconut products, resulting in inefficiencies and higher costs. Retailers and consumers expressed concerns about the lack of transparency in the supply chain, which undermines their confidence in the quality and sustainability of the products they purchase.

These perspectives underscore the need for a more integrated and transparent supply chain management system that addresses the specific concerns of each stakeholder group. Feedback also showed a strong willingness among stakeholders to adopt new technologies, provided they are easy to use and provide clear advantages in terms of efficiency, transparency, and product quality.

Quantitative data from surveys conducted among various stakeholder groups highlighted several critical trends in the coconut industry supply chain. A key finding was widespread inefficiency in current processes, with over 60% of respondents citing issues such as delayed deliveries, inconsistent product quality, and lack of real-time data. Additionally, over 70% of respondents expressed a lack of confidence in existing traceability systems, which often rely on manual records and are prone to errors.

The survey also revealed a strong interest in adopting blockchain and IoT technologies to improve supply chain transparency and efficiency. Over 80% of respondents believed that blockchain technology could significantly improve traceability, while 75% indicated that IoT could offer valuable real-time data to optimize their operations.

The results of this research illustrate the potential of modern technological solutions, especially blockchain and IoT, to greatly improve the sustainability and efficiency of the Sri Lankan coconut industry supply chain.

D. Implementation

Following analysis, the CocoGuard SCM system is implemented at designated pilot sites within the Sri Lankan coconut industry. This pilot phase allows for testing and adjustments under real-world system conditions. The effectiveness of the system is assessed through continuous monitoring and stakeholder

feedback, allowing adjustments to be made to improve performance and ensure stakeholder satisfaction.

The CocoGuard system promotes sustainability by optimizing the flow of data from farmers to consumers. Farmers use RFID and IoT sensors to monitor crop conditions and soil health, then share it with manufacturers for quality control. Data from manufacturers, combined with GPS tracking, helps distributors manage supplies and inventory more effectively. Retailers use weight sensors and RFID for accurate inventory management and to maintain product quality. Blockchain technology ensures transparency and traceability throughout the supply chain. CocoGuard improves supply chain operations by fostering sustainable practices in agriculture by reducing waste, maximizing resource efficiency, and supporting environmental compliance.

V. CONCLUSION

As explored in this research, integrating information systems and blockchain technology in Sri Lanka's coconut industry has proven to be a transformative approach to improving supply chain management practices. Using a mixed-method approach that included case studies and surveys within Sri Lanka's coconut triangle, this study has identified key areas where traditional supply chain management fails, particularly in transparency, efficiency and sustainability.

The findings of this research underscore the significant benefits of implementing smart contracts and IoT sensors within the supply chain. Smart contracts, by automating and securing transactions, reduce the need for intermediaries and mitigate the risk of fraud and errors. The use of IoT devices, such as RFID tags and sensors, provides real-time data on various supply chain parameters, including temperature, humidity and product movement. This real-time data is crucial for making informed decisions, optimizing resource use and ensuring the quality and safety of coconut products.

The proposed CocoGuard system, developed as part of this research, offers a scalable and innovative solution that addresses these challenges. By integrating smart contracts and IoT technologies, CocoGuard not only improves supply chain transparency and efficiency, but also contributes to the sustainability of the coconut industry by reducing waste, improving product traceability, and ensuring compliance with environmental regulations.

Furthermore, the insights gained from this study have broader implications that go beyond the coconut industry. The principles and technologies applied here can be adapted to other agricultural sectors, offering a model for creating more resilient and sustainable supply chains. As global demand for sustainable agricultural practices increases, systems such as CocoGuard can play a vital role in meeting these demands, thereby contributing to the overall sustainability of the agricultural sector in Sri Lanka and beyond. This research will address all the challenges encountered throughout the entire supply chain in the coconut industry.

VI. FUTURE RECOMMENDATION

At present, dashboards for farmers and manufacturers have been developed, marking a key milestone in the progress of the system. The next step will involve the creation of dashboards as well as the other features for the other stakeholders, including distributors, retailers, and consumers, ensuring that the unique needs of each role in the supply chain are addressed. Once all the dashboards and other features are complete, the entire system will be integrated and finalized for thorough testing. This testing phase will be essential to identify and resolve any issues, ensuring smooth operation and an optimized user experience. After successful testing, the system will be ready for deployment, allowing stakeholders to access real-time data and make informed decisions. Continuous monitoring and user feedback will further refine the system, improving its efficiency and reliability over time. This phased approach to development and testing will ensure the creation of a robust and effective supply chain management system.

REFERENCES

[1] M. D. Borah, V. B. Naik, R. Patgiri, A. Bhargav, B. Phukan, and S. G. M. Basani, *Supply Chain Management in Agriculture Using Blockchain and IoT*, vol. 60, no. September 2019. Springer Singapore, 2020. doi: 10.1007/978-981-13-8775-3_11.

[2] K. H. THEN, "the Current Scenario and Development of the Coconut Industry," *Plant.*, vol. 94, no. 1108, 2018, doi: 10.56333/tp.2018.010.

[3] S. M. M. Samarakoon, L. H. P. Gunaratne, and H. L. J. Weerahewa, "Determinants of Coconut Production in Large Scale Coconut Plantations in Sri Lanka: A Quantile Regression Approach," *Sri Lankan J. Agric. Econ.*, vol. 21, no. 1, p. 1, 2020, doi: 10.4038/sjae.v21i1.4641.

[4] M. A. Haq, N. R. Khan, R. Parkash, and A. Jabeen, "Impact of JIT, waste minimization, and flow management on operational performance of manufacturing companies," *Qual. - Access to Success*, vol. 17, no. 153, pp. 48–52, 2016.

[5] T. Beldek, A. K. Konyalıoğlu, and H. C. Akdağ, "Supply Chain Management in Healthcare: A Literature Review," *Lect. Notes Mech. Eng.*, no. January 2020, pp. 570–579, 2020, doi: 10.1007/978-3-030-31343-2_50.

[6] M. H. Uddin, M. R. Razzak, and A. A. Rahman, "Sustainable supply chain management practices, dynamic capabilities and competitive advantage: Evidence from Bangladesh ready-made garments industry," *Bus. Strateg. Dev.*, vol. 6, no. 2, pp. 176–188, 2023, doi: 10.1002/bsd2.232.

[7] M. R. I. slam, M. E. I. Monjur, and T. Akon, "Supply Chain Management and Logistics: How Important Interconnection Is for Business Success," *Open J. Bus. Manag.*, vol. 11, no. 05, pp. 2505–2524, 2023, doi: 10.4236/ojbm.2023.115139.

[8] N. Zarbakhshnia and A. Karimi, "Enhancing third-party logistics providers partnerships: An approach through the D.L.A.R.C.S supply chain paradigm," *Resour. Conserv. Recycl.*, vol. 202, no. January, p. 107406, 2024, doi: 10.1016/j.resconrec.2023.107406.

[9] S. S. Kamble, A. Gunasekaran, and S. A. Gawankar, "Achieving sustainable performance in a data-driven agriculture supply chain: A review for research and applications," *Int. J. Prod. Econ.*, vol. 219, no. July 2018, pp. 179–194, 2020, doi: 10.1016/j.ijpe.2019.05.022.

[10] H. O. Abdullahi, I. H. Mohamud, A. O. M. Gele, and A. Kafi, "The Role of Technology in Transforming Agricultural Supply Chain Management: Systematic Literature Review," *J. Logist. Informatics Serv. Sci.*, vol. 11, no. 1, pp. 239–251, 2024, doi: 10.33168/JLISS.2024.0116.

[11] C. Khandelwal, M. Singhal, G. Gaurav, G. S. Dangayach, and M. L. Meena, "Agriculture Supply Chain Management: A Review (2010-2020)," *Mater. Today Proc.*, vol. 47, no. xxxx, pp. 3144–3153, 2021, doi: 10.1016/j.matpr.2021.06.193.

[12] M. Alobid, S. Abujudeh, and I. Szűcs, "The Role of Blockchain in Revolutionizing the Agricultural Sector," *Sustain.*, vol. 14, no. 7, pp. 1–15, 2022, doi: 10.3390/su14074313.

[13] A. Raffington and A. Joseph, "A Qualitative Investigation into Agricultural Industry Supply Chain Efficiency During the COVID-19 Pandemic," *Researchgate.Net*, vol. 10, no. March, 2024, doi: 10.57181/ijcrst/vol10i03/76.

[14] S. Joshi, R. K. Singh, and M. Sharma, "Sustainable Agri-food Supply Chain Practices: Few Empirical Evidences from a Developing Economy," *Glob. Bus. Rev.*, vol. 24, no. 3, pp. 451–474, 2023, doi: 10.1177/0972150920907014.

[15] J. J. Rojas-Reyes, L. Rivera-Cadavid, and D. L. Peña-Orozco, "Disruptions in the food supply chain: A literature review," *Heliyon*, vol. 10, no. 14, 2024, doi: 10.1016/j.heliyon.2024.e34730.

[16] T. Alam, "Blockchain-Based Internet of Things: Review, Current Trends, Applications, and Future Challenges," *Computers*, vol. 12, no. 1, 2023, doi: 10.3390/computers12010006.

[17] D. Kumar and R. K. Dwivedi, "Blockchain and IoT Based Smart Agriculture and Food Supply Chain System," *Proc. Int. Conf. Intell. Innov. Technol. Comput. Electr. Electron. ICIITCEE 2023*, no. January, pp. 755–761, 2023, doi: 10.1109/IITCEE57236.2023.10090897.

[18] S. K. M, M. V, P. J, P. M, and R. Jothikumar, "Applying Blockchain in Agriculture: A Study on Blockchain Technology, Benefits, and Challenges," *EAI/Springer Innov. Commun. Comput.*, no. November, pp. 167–181, 2021, doi: 10.1007/978-3-030-60265-9_11.

[19] I. Hasan, M. M. Habib, Z. Mohamed, and V. Tewari, "Integrated Agri-Food Supply Chain Model: An Application of IoT and Blockchain," *Am. J. Ind. Bus. Manag.*, vol. 13, no. 02, pp. 29–45, 2023, doi: 10.4236/ajibm.2023.132003.

[20] M. B. S. M. T. Jayawardhana and B. N. F. Warnakulasooriya, "Impact of Problems associated with Supply Chain Management Practices of Wholesalers on their Business Performance in the Coconut Industry in Sri Lanka with special reference to Kurunegala District," *J. Bus. Res. Insights (former Vidyodaya J. Manag.*, vol. 06, no. 1, pp. 59–87, 2020, [Online]. Available: https://journals.sjp.ac.lk/index.php/vjm/article/view/4573

[21] S. Mandela, R. N. V. J. Mohan, and M. C. Naik, "Blockchain-based consensus for a secure smart agriculture supply chain," *Eur. Chem. Bull.*, vol. 12, no. (Special Issue 4), pp. 8669–8678, 2023.

Rubber Care - A Mobile Application to Detect The Diseases of Rubber Cultivation in Sri Lanka

Sewmini Asmadale
NSBM Green University,
Sri Lanka
sewminiasmadale@gmail.com

Mohamed Shafraz
NSBM Green University,
Sri Lanka
shafraz@nsbm.ac.lk

***Abstract* - Sri Lanka's rubber plantations are a cornerstone of its agricultural sector, but they are susceptible to various diseases that can threaten rubber production. This research presents a prototype system for the detection of rubber plant diseases, specifically designed for Sri Lankan plantations, utilizing image-based analysis without reliance on complex machine learning or convolutional neural networks. The prototype system employs a straightforward image capture mechanism, allowing local farmers and plantation workers to easily capture images of rubber plant leaves using common mobile devices. These images were used to establish visual criteria and rules for identifying and classifying different disease symptoms. The system's performance was evaluated through field testing in Sri Lankan rubber plantations, considering factors such as specificity, sensitivity, and practicality for users. The results of the prototype system demonstrate its potential for accurate and accessible identification of rubber plant diseases, providing a cost - effective solution for early disease detection and intervention in Sri Lankan rubber plantations. This research contributes to improving disease management practices in Sri Lankan rubber plantations, ensuring the resilience of this vital industry without the need for complex technology.**

Keywords—Detection, Rubber cultivation, Planters, Supervisors, image data set, prototype, mobile application

I. INTRODUCTION

Sri Lanka's rubber industry, with a rich history dating back to the early 20th century, plays a crucial role in the country's economy. The nation's favorable climate and fertile soil provide an ideal environment for producing high-quality natural rubber. However, the industry faces significant challenges from various diseases that threaten the health and productivity of rubber trees, potentially jeopardizing its sustainability. The main diseases affecting rubber cultivation includes, Pestalotiopsis, phytophthora, Corynespora, Colletotrichum, Oidium [1]. According to the reports of Rubber Research Institute of Sri Lanka these diseases not only affect the growth of rubber trees but also hinder the overall development of the natural rubber sector by limiting production. To boost natural rubber yield, consulting experts and implementing effective measures is crucial. However, the current manual diagnosis process is slow and often inaccurate, compounded by plantation workers' limited knowledge of rubber diseases. Rubber cultivation in Sri Lanka primarily takes place in several main areas across the country and Sabaragamuwa, Southern, Northwestern are the most popular rubber planting areas. [1]. considering the technology facilities and people's literacy of Sri Lankan culture in these fields, we must select the most suitable platforms for these discussions. According to these facts, physical interviews and hard copy questionaries are the most suitable platforms for the planters and we can use google forms with the facilities that they have in the situation.

The primary objective of this research work is to develop a prototype aimed at addressing the challenges faced by the rubber cultivation industry in Sri Lanka. In this undertaking, mobile applications emerge as the most suitable approach, particularly considering the needs of rubber planters in the country. By utilizing images, they can effortlessly input data into the system and obtain the desired information. Image processing encompasses a broad array of applications, such as image analysis, image classification, object recognition, and 3D imaging, among others. This project involves scanning and analyzing plant leaf images using an image classification model.[3].

II. LITERATURE REVEIEW

A. Domain Overview

The Rubber Research Institute (RRI) plays a key role in guiding the industry, emphasizing precise rubber tree cultivation, and facilitating replanting efforts. The industry spans large estates managed by reputable companies and numerous smallholders, with foreign producers capitalizing on Sri Lanka's high-quality rubber supply. Sri Lanka's resilient rubber industry involves both upstream and downstream operations. Upstream activities include the production and marketing of high-quality natural raw rubber products, such as tires, clothing, and rubber sheets. The industry consumes both locally produced and imported rubber, with notable exports of sole crepe and latex crepe rubbers to the USA, Germany, Italy, Belgium, and the UK. Factors affecting the industry include tree

age, tapping technique, and management practices. Disease control, fertilization, and pruning are crucial for maintaining healthy plantations. Leaf diseases can negatively impact rubber tree health and productivity, affecting photosynthesis and nutrient absorption.

Recent technological advancements in agricultural technology aim to enhance rubber yield by detecting and managing diseases early. State-of-the-art monitoring systems use technologies like hyperspectral imaging, drones with multispectral cameras, and IoT sensors to collect high-resolution data on parameters such as chlorophyll content, temperature, humidity, and soil moisture. AI-driven platforms analyze this data to identify disease patterns, assess plant health, and enable proactive plantation management. Predictive models help anticipate disease risks, providing recommendations for planting schedules, irrigation practices, and crop rotations, ensuring long-term sustainability, and safeguarding rubber yields.

B. Framework analysis

The development of disease detection systems in image processing using deep learning relies on key frameworks and theories for accurate and effective results. Central to many algorithms is the use of convolutional neural networks (CNNs), specialized in processing images to automatically identify illness patterns by learning from input photos. Transfer learning is crucial, involving the adaptation of pre-trained CNN models optimized for large-scale datasets like ImageNet for disease detection.[4] This approach leverages features learned from generic images, tailoring them to the specific task of illness diagnosis and often improving performance with limited data.

To enhance model resilience, data augmentation is employed, artificially expanding the training dataset through transformations such as rotations, flips, and scaling. Ensemble learning strategies, combining predictions from multiple models, contribute to improved generalization and overall accuracy.[10] In medical applications, interpretability and explainability are significant concerns addressed by methods like Grad-CAM, which offer insights into model decision-making.

C. Existing Systems

There are some existing applications that develop by using image processing techniques, and to develop this application, they use different framework designs and
algorithms. Plantrix and Agrio are applications that developed plant disease detection, and these are mostly rating applications that we can find in the Google Play Store. And Rubber Buddy is one of the applications that we can identify as developing in Sri Lanka.

Rubber Buddy is a mobile application to empower rubber planters in Sri Lanka. The application consists of
Four parts, the detection of pests in immature rubber plantations and rubber nurseries, the detection of leaf disease, the detection of cover crops, and the detection of weeds. Machine learning models created using various convolutional neural network (CNN) architectures, such as mobile net version 2 (MobileNet v2), VGG 16, VGG 19, and residual networks (ResNet), are used to recognize images captured using mobile phone cameras. Following image recognition, the program will offer rubber planters professional advice and management techniques. The program was created to function in offline mode utilizing TensorFlow Lite technology because most rubber plantations are situated in regions with poor network access. [5].

Plantix is a popular mobile application that uses artificial intelligence and image recognition to detect plant diseases and nutrient deficiencies. It covers a wide range of crops and provides recommendations for disease management and treatment. Plantix is a mobile application that utilizes image processing and computer vision algorithms to detect plant diseases and nutrient deficiencies. It employs the OpenCV framework for image manipulation and feature extraction from leaf images.[2] The application uses a vast database of plant disease images to train its machine-learning models for disease classification. Plantix incorporates a backend web development framework, such as Django or Flask, to handle user interactions and data management, allowing users to upload images, receive diagnoses, and access disease management recommendations.Agrio Plant Disease Detection Mobile Applications, Agrio is another AI-powered mobile application designed for disease detection in plants. It uses computer vision algorithms to analyze leaf images and diagnose various diseases, enabling farmers to take timely action. It is another AI-powered mobile application focused on plant disease detection. It employs computer vision and deep learning frameworks like TensorFlow or PyTorch to analyze leaf images and identify disease symptoms accurately.[8] Agrio's design involves a user-friendly interface that allows farmers to capture images of affected leaves using their smartphones. The application's machine learning models are continuously updated with new data to improve disease detection accuracy.

Plant diseases are crucial to the development of agriculture since they can particularly affect the quality and quantity of plants. In general, fungi, bacteria, viruses, and mold are the causes of plant illnesses. Plant diseases are often diagnosed by farmers or professionals using only their eyes. However, this method can be time-consuming, expensive, and inaccurate. As a result, deep learning-based plant disease detection and categorization offer a quick and accurate way. For the purposes of study, instruction, and analysis, photographic images of plant infection signs are used to diagnose plant diseases.[6] Deep learning technology and computer image processing are used to provide speedy and accurate detection, and studies show that deep learning

techniques are useful for classifying plant diseases.[7]. Enhancing image analysis's dependability, correctness, and accuracy for spotting and classifying plant sickness has been a top priority. To extract plant disease representation from trained CNNs, using visualization approaches.

III. METHODOLOGY

The goal of Design Science Research (DSR), a research methodology, is to develop novel artifacts that solve real-world issues in a variety of fields. It is especially pertinent in industries like information systems, engineering, and technology were developing new solutions and designing them is crucial. DSR seeks to develop physical, practical objects that can be used in real-world settings in addition to new information.

The research focuses on addressing challenges in detecting diseases in rubber plant leaves through image processing. Motivated by the critical need for swift and accurate identification, the process involves a literature review, theoretical foundation, conceptual design, and development. The system is implemented, tested, and empirically evaluated for accuracy and robustness using diverse datasets. An iterative improvement phase refines the system based on evaluation results, ensuring continuous enhancement. Comprehensive documentation preserves methodologies and outcomes, and communication and knowledge dissemination share research outcomes with peers, practitioners, and stakeholders through conferences and academic journals.

A. Fact collection mechanisms

In this research project focused on developing an efficient method for detecting leaf disease on rubber plants, questionnaires and interviews play pivotal roles in gathering valuable insights from rubber producers and experts. The questionnaires, available in Sinhala, Tamil, and English, aim to inclusively capture practitioners' experiences, providing crucial information on disease frequency, challenges faced, and current management strategies. The goal is not just to collect data but to emphasize the necessity for a technologically advanced disease detection system. The gathered information serves as a compelling argument for the development of a rubber plant leaf disease detection system based on image processing. It positions the rubber farmers as key contributors to advancing disease control paradigms by bridging the gap between real-world knowledge and technological advancements.

Interviews with rubber experts and supervisors are carefully planned to extract in-depth knowledge and extensive experience in rubber farming and disease management. These interviews delve into complex insights beyond surface observations, aiming to understand various facets of disease detection, treatment, and prevention. The experts' opinions on disease characteristics, challenges unique to rubber production, and their decision-making processes are explored. Overall, the combination of questionnaires and interviews creates a comprehensive foundation for the development of a robust rubber plant leaf disease detection system, aligning technology with the practical experiences and expertise of those involved in rubber cultivation.

B. Data Analysis

In this research project, thematic analysis was employed as the selected method for analyzing the gathered data. The data collection process encompassed two separate approaches,

- Conduct interviews and questionnaires for gathering data to identify the requirements.
- Collecting images for training the model.

The first part of thematic analysis was done by using research questions and answers. According to the research question there are main 3 objectives in this work.

- Object 1 - To identify the most prevalent diseases in rubber cultivation in Sri Lanka.
- Object 2 - To identify the existing methods of disease detection and challenges.
- Object 3 - To propose a user-friendly solution suitable for the rubber industry in Sri Lanka.

In first objective, analysis aimed to identify the most common diseases affecting rubber plants in Sri Lanka. The data collected were analyzed to investigate the types of diseases that rubber farmers and agricultural officers had observed or dealt with. Additionally, the analysis examined the frequency of disease occurrences and the methods used for reporting diseases. The second objective was focused on understanding the current disease detection methods employed in rubber cultivation. The various approaches, including visual inspection by farmers, laboratory-based testing, and the use of government or non-government support, were explored. Furthermore, the analysis delved into the challenges associated with these methods, such as time-consuming processes, accuracy, and accessibility of support.

Final objective was aimed to propose a user-friendly solution for disease detection in the Sri Lankan rubber industry. The data were analyzed to understand preferences for mobile apps or handheld devices and the desired features of such a system, including real-time analysis and a user-friendly interface. Additionally, the analysis examined respondents' awareness of and willingness to adopt technology, along with any concerns or reservations they may have had about technology adoption and access to technology in rubber cultivation in Sri Lanka.

In the data set creation process, there are several steps to follow. The image dataset for the rubber tree plants leaf disease detection system was developed by capturing images of leaves affected by the above selected disease types. The research objectives were clearly defined to focus on the accurate identification of diseases in rubber tree leaves. Image collection involved methods, such as capturing photos using cameras and smartphones, along with augmenting existing datasets through transformations like rotation,

scaling, and cropping to increase image diversity. All images are captured by using white background and image size was 1936 x 129.

Images were captured, meticulous labeling was undertaken, and recording essential details for each image to provide crucial information for subsequent training and evaluation of the disease detection system. The images were organized and stored in distinct folders, categorized based on the specific disease affecting the leaves. This organizational structure was implemented to ensure the dataset's clarity and accessibility. Data augmentation techniques were applied to create variations of existing images, potentially enhancing the model's ability to generalize and accurately detect diseases.[9] The image dataset for the rubber tree plants leaf disease detection system was crafted through a systematic process that encompassed clear objective definition, diverse image collection methods, thorough labeling, organized storage, preprocessing, and strategic dataset division for model development and evaluation.

Fig 1. Healthy Rubber Leaf

Fig 2. Phytophthora

Fig 3. Pestalotiopsis

Fig 4. Pestalotiopsis

IV. PROPOSED SYSTEM

Proposed system application's home page serves as the central hub within the Research System mobile application, serving as the main screen that users encounter upon launching the app. Its primary purpose is to offer users a comprehensive overview of the app's diverse features and functionalities. Beyond being an informational gateway, the home page is designed for user convenience, enabling swift access to the most frequently performed tasks. This intuitive layout ensures that users can efficiently navigate through the app, streamlining their interactions and enhancing their overall experience. In essence, the home page functions as the app's core, providing users with both a snapshot of its capabilities and a practical launchpad for their research-related activities.

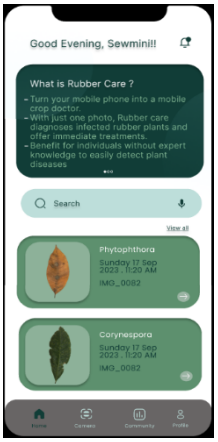


Fig 5. Landing Page

Fig 6. Home Page (a)

Once an image is captured and incorporated into the scanning process, users are presented with a dedicated disease decryption page within the Research System mobile application. This specialized page serves as a valuable resource for users seeking in-depth information about various diseases. Here, users can access comprehensive details regarding diseases, including vital information about their symptoms, causes, and available treatments. This feature empowers users with the knowledge needed to better understand the diseases they are researching, aiding in their quest for information, diagnosis, and treatment options.

The disease decryption page represents a pivotal aspect of the app, providing users with the means to access critical health-related insights quickly and efficiently. Each disease displays separate information according to the detection disease.

The diseases treatment page provides users with the immediate treatments that they can do for the diseases and recommend the fungicides they can use for the diseases separately. Users can learn more details using Read more on Wikipedia. Users can share their images with others and save the images.

After detecting the diseases users can share their images with the others in the community with adding enquiry. And by using the community page users can like and comment on the other images and answer foe their questions. Community page displays the username and the location.

The recent page within the Research System mobile application offers users a convenient and efficient way to access their most recently detected images. This feature serves as a practical resource for users to quickly locate and retrieve resources they have recently used or are currently working on. By displaying a list of the most recently detected images, this page simplifies the process of revisiting important materials, streamlining the user experience. Whether users need to reference recent research, continue work on an ongoing project, or simply keep track of their most recent activities, the recent page provides easy access to these resources, ultimately enhancing productivity and user satisfaction within the app.

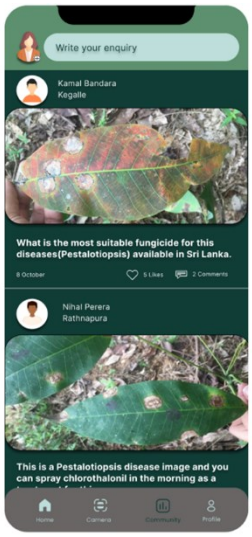


Fig 9. Community Page

Fig 10. Recent Page

V. DISCUSSION

In the Discussion, it is crucial to emphasize the significance of the results obtained in the context of the original aims and objectives of this research, which centered on the development of prototypes and an image dataset for a system to detect plant diseases in rubber cultivation. The results obtained from these prototypes and the dataset hold substantial implications for the rubber industry. The prototypes' potential to accurately detect diseases in rubber plants and the creation of a comprehensive image dataset offer a promising direction for cost-effective and efficient early disease identification. This could significantly contribute to mitigating crop losses and enhancing overall rubber cultivation practices. These findings are particularly important for the rubber industry, as they present an opportunity to advance productivity and sustainability.

The research findings should be compared to previously published data in the field. Upon such comparisons, it becomes evident that these prototypes and the dataset provide valuable resources for further research in disease detection. This comparative analysis reaffirms the novel contribution of this work in the field of plant disease detection.

VI. CONCLUSION

In conclusion, the major conclusions drawn from this investigation can be concisely stated as follows: The development of prototypes and an image dataset for a system to detect plant diseases in rubber cultivation show promise in providing practical solutions for an industry that often grapples with disease-related losses. To pave the way for future research endeavors, it is suggested to explore the integration of machine learning models to enhance disease prediction and the implementation of real-time monitoring systems for early disease detection, utilizing the resources provided by these prototypes and the dataset. These future research avenues hold the potential to further advance disease management in rubber cultivation and improve the industry's overall sustainability and productivity.

REFERENCES

[1] Sri Lanka Export and Development Board, 2023. Natural Rubber Industry In Sri Lanka - EDB Sri Lanka. [Online] Available at: https://www.srilankabusiness.com/rubber/about/industry-capability.html [Accessed 01 July 2023].

[2] Rupavatharam S and Kennepohl A and Kummer B and Parimi, V. (., 2018. Automated plant disease diagnosis using innovative android App (Plantix) for farmers in Indian state of Andhra Pradesh. Rupavatharam S and Kennepohl A and Kummer B and Parimi, V (2018).

[3] Sri Lanka Export and Development Board, 2023. Natural Rubber Industry In Sri Lanka - EDB Sri Lanka. [Online] Available at: https://www.srilankabusiness.com/rubber/about/industry-capability.html[Accessed 01 July 2023].

[4] Swain Anisha Image Processing Framework for Machine Learning, 2018. What is an Image Processing Framework for Machine Learning[Online]Available at:https://medium.com/@anishaswain/noise-filtering-in-digital-image-processing-d12b5266847c [Accessed 13 July 2023].

[5] Aparna Jayawardena Kasuni Ganegoda Sakuni Imbulana Gavin Gunapala Nuwan Kodagoda Thilini Jayasinghe, 2022. Rubber Buddy: A Mobile Application to Empower Rubber Planters of Sri Lanka. Aparna Jayawardena Kasuni Ganegoda Sakuni Imbulana Gavin Gunapala Nuwan Kodagoda Thilini Jayasinghe,, Issue 09-10 December 2022.

[6] Ayesha Siddiqua Muhammad Asha Kabir Tanzina Ferdous Israt Bintea Ali, 2022. Evaluating Plant Disease Detection Mobile Applications: Quality and Limitations. Ayesha Siddiqua Muhammad Asha Kabir Tanzina Ferdous Israt Bintea Ali,.

[7] Laxmi M. Tech Scholar, 2022. Detection And Classification Of Plant Leaf Disease Using. [Online] Available at: https://www.ilkogretim-online.org/[Accessed10072023].

[8] Zeng T, (28 February 2022. Rubber Leaf Disease Recognition Based on Improved Deep Convolutional Neural Networks With a Cross-Scale Attention Mechanism. [Accessed10JULY2023].

[9] J. Arun Pandian, G. Geetharamani and B. Annette, "Data Augmentation on Plant Leaf Disease Image Dataset Using Image Manipulation and Deep Learning Techniques," *2019 IEEE 9th International Conference on Advanced Computing (IACC)*, Tiruchirappalli, India, 2019, pp. 199-204. [Accessed 07 July 2023].

[10] M. Kirola, K. Joshi, S. Chaudhary, N. Singh, H. Anandaram and A. Gupta, "Plants Diseases Prediction Framework: A Image-Based System Using Deep Learning," *2022 IEEE World Conference on Applied Intelligence and Computing (AIC)*, Sonbhadra, India, 2022, pp. 307-313, [Accessed 10 July 2023].

Digital Intervention for Managing Stress of Undergraduates

Shimary Fernando
Faculty of Computing
NSBM Green University
Homagama, Sri Lanka
shimaryfernando@gmail.com

Mohamed Shafraz
Faculty of Computing
NSBM Green University
Homagama, Sri Lanka
shafraz@nsbm.ac.lk

***Abstract* — Using a thorough study of the research, it was determined that many students struggle with stress during their university years, which in turn tends to affect their overall academic performance. The relationship between stress and diminished academic achievement was clearly shown in prior research. The identified stress primarily fell into two distinct categories: Academic Stress and Environmental Stress. This study was conducted to identify effective stress management techniques. The stigma currently associated with seeking counselling services, which discourages students from engaging in it, appeared as a major barrier. To gather data, interviews were conducted with two experienced counselors, while questionnaires were distributed among undergraduates. The three objectives of the study and the theoretical framework outlined in the literature review guided the careful design of both sets of questions. Then Thematic Analysis was conducted, and 6 Themes and 16 sub themes were identified. Simultaneously, to clarify whether prototype or questionnaire is reviewed. Based on the findings from Interviews and Questionnaire, both functional and nonfunctional requirements were gathered. According to the gathered requirements system was designed, considering essential UI/UX principles for an optimal user experience.**

Keywords—stigma, stressors, stress, academic performance, undergraduates, web application

I. INTRODUCTION

Mental health is important because it becomes a part of a human life, and it impacts on the overall health of the person including physical, mental, and social well-being. It is not just about being sick and showing diseases [1]. The mental health services are not received the same for each person. In high-income countries, this percentage is 70%, but in low-income countries, it drops to only 12% receiving care services [2]. So, the Importance is varied globally and concerns regarding mental health are different country to country.

In Sri Lanka has an average level of economic development since it is considered a country with middle-income levels. For many years, Sri Lanka has been able to produce positive results in mental health. Also, it is very associated with the family unit [3].

According to previous research, 12% - 50% of undergraduates have one or more mental disorders [4]. On campuses, mental health issues have grown significantly more important. Undergraduates who are struggling with their mental health may also have concerns about their physical health and their academic performance. Particularly depression and anxiety are linked with lower GPAs, more risky behaviors, dropping out of college, and even suicidal ideation. Universities must raise awareness of mental health issues, decrease stigma, offer accessible support services, and put preventive measures in place to address these issues. Universities may support students' academic and personal success by placing a high priority on mental health [5].

High expectations and demands have been placed on undergraduates due to the rapid advancements in technology, education, culture, and society. High expectations and demands, often unclear and unrealistic within socio-cultural, economic, and regulatory contexts, contribute to significantly higher stress levels compared to previous generations [6].

Previous research indicates that undergraduates experiencing high levels of psychological distress may struggle with their academics and face difficulties completing their studies. There is a notable correlation between psychological distress and poor academic performance. Nearly 20% of undergraduates reported missing classes due to mental health issues, and almost 44% experienced a decline in academic performance during their undergraduate studies because of mental health challenges [7].

The research targets undergraduates experiencing stress during their studies. Over the past decade, there has been a troubling increase in anxiety and stress among students, with 31% of undergraduates worldwide reporting mental health issues in the last year. This statistic reflects a significant rise in mental health challenges among undergraduates [8]. The research also explores how stress affects academic performance and investigates IT interventions designed to manage stress and improve university students' academic outcomes.

A. Stress among University Students and Factors Affecting the Stress

Previous research has been conducted on Stress in university students with the focused of academic stress. Academic stress is the term for pressures that arise from being involved in organizations, working

on group projects, and taking on much university work [9]. University-level academic stress can be brought on by packed lecture rooms, the semester system, and a lack of tools to complete academic work. The academic setting is particularly stressful because of the time limits and pressure to do well on tests or exams [10].

Previous research has mentioned that environmental stress can also affect the University Students. There are so many factors that can affect the environmental stress of the student. Such as peer relationships, instructor treatment, diversity, support groups, community integration, in-class engagement, and contacts with faculty, home sickness, relationship issues, and financial issues [10].

B. Research Problem

The study is mainly targeted at undergraduate students, as described in the above section. According to past research, the university years represent an important developmental phase in the lives of individuals. This period is characterized by the transition from late adolescence to emerging adulthood, a stage of life that encompasses the late teens and early twenties [11].

Research, including studies by the American College Health Association (2018), shows that university students experience higher stress levels and mental health issues compared to the general population [12].

As a solution, most of the universities have their own counsellor to solve their problems and share their problems. So, at some sort of level, stress can be managed. But it's challenging for students and counselors as well. It takes time to the treatment. Back in 2018, according to the National College Health Assessment, at a point 86% of students feel overwhelmed, 69% of the students felt sad, 65% of the students feel loneliness and 55% of the students felt hopeless, 13% of students feel suicidal. Furthermore, that research indicated severity of mental health and treatment seeking students [8].

With the technological revolution, counselling services are provided online by developed countries. So, the E-therapist concept has been popular for years. However, there is a problem with what type of E- therapist is wanted by a particular university student.

The proposed system aims to enhance counseling services by tracking students' stress levels based on their degree program and academic year. It provides tailored coping strategies for stress management, with suggestions categorized according to the specific stress factors the student is experiencing. By implementing this system, the goal is to improve undergraduate GPA outcomes.

C. Research Objectives

The research was conducted based on three main objectives.

1. To identify Academic and Environmental factors for stress of university students which affect their academics.
2. To evaluate and study existing IT solutions for stress management.
3. To design the best IT solution for managing stress to improve their academics.

D. Based Theory about Stress regarding the Current Study

There are many theories about stress, each of which presents a different viewpoint on its causes and effects. In the context of this study, Lazarus's cognitive appraisal theory was selected as the foundational framework. This is a theory by Lazarus that focuses on how people perceive and cope with stressors. It suggests that how you interpret a situation (appraisal) influences your emotional response and coping strategies. A previous study looked at how Lazarus' theory applies to university students' performance. Students who saw an upcoming exam as a threat were more likely to experience negative emotions and use coping methods centered on managing those emotions rather than directly addressing the stressor [13].

II. RESEARCH METHODOLOGY

A. Research Paradigm

The research will follow both positivism and interpretivism. The positivism approach is being used to collect quantitative data and interpretivism approach is being used to collect qualitative data. Interviews will be conducted with the counsellors, along with a questionnaire to gather both types of data. The questionnaire will be distributed among students from private and government universities.

B. Research Approach

The research was initiated by observing issues that were brought up in the context of the health industry. The theory in earlier stress-related literature provides substantial support for this observation. The study aims to address similar problems faced by Sri Lankan university students. Given the nature of the research, a deductive reasoning strategy is applied but in the present research, some of the steps were altered.

C. Research Strategy

Research is conducted by using action research to address the problem of reduced academic performance due to academic and environmental stress among students. The solution involves developing a web application designed to help manage stress, which will be implemented for university undergraduates. Literature review findings, which confirm the negative impact of stress on

academic performance, will serve as a benchmark to assess changes after the intervention.

III. Results and Evaluation

A. Data Collection and Data Analysis

1) Thematic Analysis

The data which is gathered through Interviews is analyzed by using Thematic analysis. This method helps to systematically organize and interpret the qualitative data to identify recurring patterns or themes within the collected information through Interviews. As a result of this analysis, 6 Themes and 16 Sub Themes could be identified.

TABLE 1. Themes and Sub Themes

Themes	Sub Themes
Involvement Counselors for support of students	Existing Communication approaches between undergraduates and counselors
	Existing Communication Approaches Associated Challenges
	Requirement of Counselling
Involvement of stressors of students	Academic Stressors of Undergraduates
	Environmental Stressors of Undergraduates
	Other Factors which affect the Undergraduates
	Links between stress levels and academic achievement
Emotional Responses and Appraisal	Primary Appraisal Methods of Undergraduates
	Secondary Appraisal Methods of Undergraduates
Coping Mechanism	Coping Strategies for the Academic Stressors
	Coping Strategies for the Environmental Stressors
	Types of Coping
Identify Individual Differences	Personal Traits
Use of IT Solutions for Stress Management	Existing IT solutions
	Limitations and Considerations
	Solutions
	Desired Features

2) Mapping data collection of questionnaire with results review

The student survey reveals significant academic pressures among respondents, with 58.9% experiencing high stress from academic demands and 27.4% reporting moderate stress. Environmental factors like living conditions also contribute to high stress levels for 64.4% of students. Stress appears to negatively impact academic performance, as shown by GPA distributions: 43.8% have a GPA of 3.0-3.29, while 32.9% have a GPA of 2.0-2.9. Despite 64.4% knowing about university counselling services, many are unfamiliar with them, potentially contributing to stress. Almost all undergraduates lack knowledge of web-based stress management tools, but a majority are open to using one integrated into the university platform, with 54.8% being very likely to use it.

B. Functional Requirements and Non Functional Requirements

A web application was properly created, incorporating both functional and non-functional requirements resulting from the themes that came up, based on these findings. The sample population's current stress management strategies are considered in this application, which also discusses the difficulties these strategies face.

1) Functional Requirements

- Signup to the System
- Student Login to the System
- Access to the Dashboard
- Evaluate the Stress Level according to the PSS Score and activities are suggested according to the Score.
- Self-Evaluation – Select category/ categories of stressors.
- Self-Evaluation – Select type of stressors.
- Self-Evaluation – Activity Suggestions.
- Instructions and Tutorial regarding suggested activity can be viewed accordingly.
- The student should be able to chat with the Counsellor of the University. Also, Student should be able to schedule a meeting by clicking on Schedule.
- Admin should be able to add activities regarding stressor by accessing the different modules in the system.

2) Justification of Functional Requirements

While doing Interviews, their main concern was to suggest coping strategies according to their stress level. Also, another main concern was to evaluate their stress level on their own. It became clear that they wanted more than just advice on how to effectively manage their stress; they also wanted the capacity to independently evaluate and understand their own stress levels. By considering this requirement the undergraduate should be able to evaluate their stress level according to the PSS Score and Activities are suggested according to their stress level. This gives undergraduates the ability to assess their own levels of stress using a valid metric, providing them with an accurate representation of their present mental health. Also, the system facilitated personalized activity suggestions according to their stressors under the main two categories. According to the terminology used by counselors, these activities essentially equal coping mechanisms. This approach tries to fill the gap between expert terminology and familiar language. The system follows Emotion – Focused Coping which was mentioned in the interview as Coping Type. So, the

stressors and coping strategies which are suggested by the system are aligned with the mentioned coping strategies and stressors in the interview by counselors. Another mentioned point in the interview was that the student may have a combination of stressors. To accommodate this, students can select and address multiple stressors simultaneously. The coping strategies offered are comprehensive and specifically adapted to each user's unique stress profile because of this subtle methodology. In a nutshell, the system seeks to encourage undergraduates to take the initiative in their mental health by encouraging empowerment and self-awareness along with practical answers.

3) Non-Functional Requirements

- To secure the user's personal data, security implementation should be there.
- Since this is a stress management system, the user interface should be more interactive.
- The system should be able to generate audits. (Logs)
- The system should be able to respond promptly.
- The system should be accessible to many students at once.

4) Justification of Non - Functional Requirements

Data privacy and confidentiality were the primary concerns for the students. To address this, security measures like two-factor and multi-factor authentication are implemented, along with role-based permissions for all modules. The user interface is designed to enhance user experience and ensure ease of use, particularly for students dealing with stress.

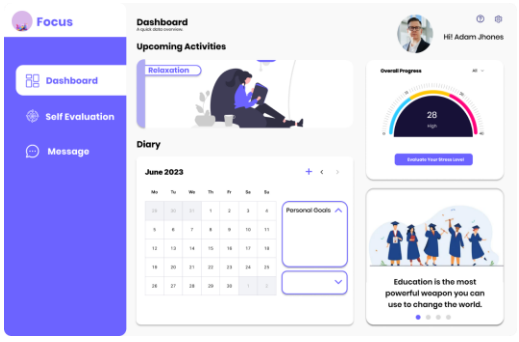


Fig. 1. Student Dashboard

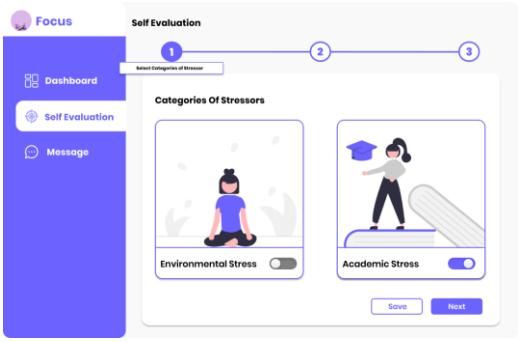


Fig. 2. Self-Evaluation (Step 01)

Fig. 3. Self-Evaluation (Step 02)

C. Result on Design Evaluation

After completing the UI design, a prototype was presented to selected students and a counsellor. The students were satisfied with the self-evaluation and stress measurement features. The counsellor was also satisfied but suggested adding a customizable theme color option, which was included in a new "setup" module where students can choose primary and secondary colors. Both the chat feature and meeting scheduling options were well-received. The counsellor noted that the app wouldn't conflict with existing university counselling services and could enhance student engagement with these services. Additional student suggestions, like time management tools and priority lists, will be considered in the project's second phase. Overall, feedback was positive, with both favorable and critical input informing future development to help reduce student stress.

IV. Conclusion

The research project provided a deeper understanding of stress management among undergraduates, enhancing both my academic performance and my ability to engage in discussions on mental health and well-being. The process also developed my analytical and critical thinking skills, particularly in analyzing information from various literature reviews. I learned valuable time management and workload balancing skills, allowing me to prioritize tasks and strengthen my abilities. The

project included both rewarding and challenging learning experiences, particularly in reading and analyzing literature to identify research gaps effectively.

V. FUTURE WORKS

With limited time, the system cannot be implemented. So, the system will be implemented according to the designed UIs. Additional features will be incorporated based on the specific requirements of universities to enhance the system further. To increase the accuracy of stress level measurements and provide more effective coping strategies, AI and IoT devices will be integrated in future updates.

ACKNOWLEDGMENT

I would like to take this opportunity to express my sincere gratitude to my guiding light, Dr. Mohammed Sapraz, who was with me to the end of my final year of research, guiding me and sharing his knowledge and experience. I would also like to thank all the other lecturers who gave their continuous support throughout the creation and delivery of the final year's research. Last but not least, my heartful gratitude goes to my family members and friends for their understanding and encouragement until the conclusion of this research.

REFERENCES

[1] Pimentel, L. (2019). The Importance of Mental Health in the Mexican Public Health System. *Mexican Journal of Medical Research ICSA Biannual Publication*, *7*(13), 1–6. https://repository.uaeh.edu.mx/revistas/index.php./MJMR/issue/archive

[2] Cuijpers, P., Javed, A., & Bhui, K. (2023). The WHO World Mental Health Report: a call for action. *The British Journal of Psychiatry*, 1–3. https://doi.org/10.1192/bjp.2023.9

[3] Liyanage, U. L. N. S., Rowel, W. W. J. S. M., Hewawitharana, U. H., De Silva, V., & Dissanayake, D. M. T. (2018). Stigma due to their work in mental health, among mental health professionals in Sri Lanka. *Sri Lanka Journal of Psychiatry*, *9*(1), 4. https://doi.org/10.4038/sljpsyc.v9i1.8174

[4] Bruffaerts, R., Mortier, P., Kiekens, G., Auerbach, R. P., Cuijpers, P., Demyttenaere, K., Green, J. G., Nock, M. K., & Kessler, R. C. (2018a). Mental health problems in college freshmen: Prevalence and academic functioning. *Journal of Affective Disorders*, *225*, 97–103. https://doi.org/10.1016/j.jad.2017.07.044

[5] Oswalt, S. B., Lederer, A. M., Chestnut-Steich, K., Day, C., Halbritter, A., & Ortiz, D. (2020a). Trends in college students' mental health diagnoses and utilization of services, 2009–2015. *Journal of American College Health*, *68*(1), 41–51. https://doi.org/10.1080/07448481.2018.1515748

[6] Ramli, N. H. H., Alavi, M., Mehrinezhad, S. A., & Ahmadi, A. (2018). Academic stress and self-regulation among university students in Malaysia: Mediator role of mindfulness. *Behavioral Sciences*, *8*(1). https://doi.org/10.3390/bs8010012

[7] Sharp, J., & Theiler, S. (2018). A Review of Psychological Distress Among University Students: Pervasiveness, Implications and Potential Points of Intervention. In *International Journal for the Advancement of Counselling* (Vol. 40, Issue 3, pp. 193–212). Springer New York LLC. https://doi.org/10.1007/s10447-018-9321-7

[8] Lattie, E. G., Adkins, E. C., Winquist, N., Stiles-Shields, C., Wafford, Q. E., & Graham, A. K. (2019). Digital mental health interventions for depression, anxiety and enhancement of psychological well-being among college students: Systematic review. *Journal of Medical Internet Research*, *21*(7). https://doi.org/10.2196/12869

[9] Karaman, M. A., Lerma, E., Vela, J. C., & Watson, J. C. (2019b). Predictors of Academic Stress Among College Students. *Journal of College Counseling*, *22*(1), 41–55. https://doi.org/10.1002/jocc.12113

[10] Reddy, K. J., Menon, K. R., & Thattil, A. (2018). Academic stress and its sources among university students. *Biomedical and Pharmacology Journal*, *11*(1), 531–537. https://doi.org/10.13005/bpj/1404

[11] Harrer, M., Adam, S. H., Baumeister, H., Cuijpers, P., Karyotaki, E., Auerbach, R. P., Kessler, R. C., Bruffaerts, R., Berking, M., & Ebert, D. D. (2019). Internet interventions for mental health in university students: A systematic review and meta-analysis. *International Journal of Methods in Psychiatric Research*, *28*(2). https://doi.org/10.1002/mpr.1759

[12] Rith-Najarian, L. R., Boustani, M. M., & Chorpita, B. F. (2019). A systematic review of prevention programs targeting depression, anxiety, and stress in university students. In *Journal of Affective Disorders* (Vol. 257, pp. 568–584). Elsevier B.V. https://doi.org/10.1016/j.jad.2019.06.035

[13] Deng, Y., Cherian, J., Khan, N. U. N., Kumari, K., Sial, M. S., Comite, U., Gavurova, B., & Popp, J. (2022). Family and Academic Stress and Their Impact on Students' Depression Level and Academic Performance. *Frontiers in Psychiatry*, *13*. https://doi.org/10.3389/fpsyt.2022.869337

Support Services Mobile Application for Micromobility Users

UU Upasena and Dasuni Ganepola
Faculty of Computing
General Sir John Kotelawala Defence University
Rathmalana, Sri Lanka
38-bis-0008@kdu.ac.lk

***Abstract*— In the rapidly urbanizing landscape of Sri Lanka, micromobility has emerged as a key solution to address challenges related to traffic congestion, high fuel costs, limitations on vehicle imports and environmental degradation, while also offering significant health benefits and serving as a popular leisure activity for many. Electric bicycles (e-bikes) have become a popular trend in the Sri Lankan micromobility market as they offer the convenience of a traditional bicycle with the added benefit of an electric motor, making them an ideal choice for daily commuters. Despite its growing popularity, micromobility users face significant obstacles due to the lack of centralized information on essential support services, such as secure parking, dedicated cycling lanes, charging stations, and repair shops. This paper presents the development of a support services mobile application tailored specifically for micromobility users in Sri Lanka. The application integrates features such as real-time location tracking, navigation, and social networking to enhance the commuting experience. Leveraging technologies like Flutter, Firebase, and Google Maps Platform, the application provides a seamless and supportive platform that addresses the critical gaps in current micromobility infrastructure. Through this research, we aim to contribute to the micromobility domain by offering a solution that enhances user safety, convenience, and satisfaction, ultimately fostering a more sustainable and efficient urban transportation system.**

Keywords—micromobility, mobile application, sustainability

I. INTRODUCTION

In the rapidly evolving landscape of urban transportation, micromobility has emerged as a pivotal solution, offering an efficient and eco-friendly alternative to traditional vehicular travel. This paper presents the development of a support services mobile application tailored specifically for micromobility users. Micromobility vehicles, defined as electric or human-powered two-wheel vehicles weighing less than 350 kg and capable of speeds up to 45 kmph, have gained immense popularity globally. According to recent statistics, the micromobility industry is a $102 billion sector, with over one billion vehicles in use worldwide. Micro-mobility is an environmentally, socially, and economically sustainable transportation mode, which is affordable and easily assessable and requires little training to control. Micro-mobility operations, primarily in metropolitan areas where they serve as a practical solution to high fuel costs, increased public transport fares, and severe traffic congestion as Sri Lanka is undergoing rapid urbanization, with increasing levels of personal vehicle ownership and usage, and consequent environmental degradation, including air pollution and carbon emissions [1].

In the Sri Lankan context, micro-mobility users face significant challenges due to the fragmented and often inaccessible nature of essential information about support services. These support services, which include secure parking areas, dedicated cycling lanes, charging stations for electric bikes, and repair shops, are crucial for ensuring the safety and convenience of micro-mobility users. To ensure a safety ride micromobility users need to identify bicycle tracks, and cycle lanes on roads also to avoid the traffic congestion caused by other fossil powered vehicles. However, the lack of a centralized system to locate and access these services makes it difficult for users to fully leverage the benefits of micromobility. To ensure a safety ride micromobility users need to identify bicycle tracks, and cycle lanes on roads also to avoid the traffic congestion caused by other fossil powered vehicles. This research addresses the above gap by developing a mobile application that provides a range of support services tailored to micromobility users. The application will include user authentication and profile management, map and navigation features, information on cycle repair shops and charging stations, and social features for maintaining a network of friends. By integrating these functionalities, the application aims to provide seamless and supportive experience for micromobility users, ensuring they have access to the necessary resources for a safe and efficient commute. This research contributes to the domain by the development of a support services mobile application

for micromobility users, addressing the various challenges they face and enhancing their commuting experience.

This research paper is organized as follows: Section 2 literature review, Section 3 Design overview, Section 4 development, Section 5 Discussion and finally the Conclusion.

II. LITERATURE REVIEW

A. Overview of Micromobility

Electric or human-powered two-wheel vehicles that are less than 350 kg and have a top speed of 45 kmph, are widely known as micromobility vehicles [2]. According to statistics from the world's largest micromobility producers China, United States, and India there are over 1 billion of micromobility vehicles worldwide of 102-billion-dollar industry in 2023. According to Statistics of Department of motor traffic there are 2 million of estimated number of bicycles and micromobility vehicles are in Sri Lanka. Most of the commuters traveling within/to metro cities in Sri Lanka are using micromobility vehicles to navigate to their destinations. Due to high gas prices, increasement of fair for the public transport and traffic congestion caused by other fossil fuel-powered vehicles lead to an increase in the use of micromobility vehicles.

Cycling is considered as one of the best ways to harmonize daily life demands, mental and physical health needs, and sustainability challenges in urban scenarios. The adverse effects of relying heavily on automobiles have gained widespread acknowledgment in recent years. Consequently, cities worldwide are directing investments towards alternative transportation options that hold the promise of enhancing public health, decreasing air pollution and carbon emissions, and alleviating traffic congestion [3]. Furthermore, as a mode of transport encouraging physical activity and interaction with the surroundings, cycling has the potential to enhance both personal health and well-being. Additionally, bicycles address issues of social exclusion by being cost-effective and accessible to a wide range of individuals. To summarize, bicycles offer individual and societal advantages, contributing positively to physical health, mental well-being, environmental concerns, and inclusivity [4]. During and after the COVID-19 pandemic, all transport models decreased expect of cycling, increasement of walking and cycling have been identified as key mobility trends in urban areas. Average daily kilometers cycled increased Times of cycling trips and trip purposes changed (more leisure, more midday), length of cycling trips changed (longer trips for leisure purpose), cycling increased in more transit-oriented cities while it decreased in more bicycle-oriented university cities and more bicycle activities in rural areas for recreational purpose [5].

B. User Behaviour and Preferences

According to the survey done by Department of Town and Country Planning, University of Moratuwa, by interviewing about 200 bicycle riders 30%,30%,15%,15%,10% of population state that health effective, ability to avoiding traffic congestion, Time saving, cost effective and no required more space to parking as respectively as the reasons to use bicycles [5]. The results of the interviews conducted show that most of the people ride bicycles to avoid traffic and to assure their health. These two aspects show the emerging motivation fact behind the riders, which is a prospective aspect. It further elucidates a promising future, where many bicycles can be seen on the roads, instead of automobiles.

C. Use of Mobile Applications Among Users

The survey conducted by Cycling Weekly [6] to understand the defining traits of modern amateur cyclists, with the participation of nearly 600 cyclists found that 71% of respondents used a navigation app to travel. Nowadays most micromobility riders use navigation applications in Sri Lanka context also. The Department of Town and Country Planning at the University of Moratuwa conducted a survey gathering information from individuals regularly commuting to the Colombo Metropolitan Area. The data collection involved the distribution of questionnaire forms both in person to daily visitors and through online surveys. Additionally, discussions were held with taxi drivers and travelers. The questionnaire covered general information about the travelers, their familiarity with the area, route and destination selection criteria, reasons for using navigation apps, frequency of travel to specific destinations, and their perceptions of future travel patterns. The data collection process included both in-person surveys and online responses. The results shows that gender affects route selection of familiar area, men use navigation compared to other age brackets. Analyzing transportation modes, three-wheelers (91%) exhibit a greater tendency to use navigation apps for route selection in both familiar and unfamiliar areas compared to other modes. Additionally, for travelers with leisure as their primary purpose, there is an increased utilization of navigation apps (80%) for route selection in both familiar and unfamiliar areas. Women demonstrate a higher likelihood (26%) of using navigation apps to find their destination in familiar areas compared to men. In terms of travel mode, cars (29%) are more prone to using navigation apps to find destinations in familiar areas than other modes. For unfamiliar areas, motorbike users exhibit the highest reliance on navigation apps (97%) compared to other transportation modes. Examining the traveler's purpose, those traveling for studies tend to use navigation apps for finding destinations in both familiar and unfamiliar areas. Apps more (67%)

compared women. They are inclined to use navigation applications for route selection, even in unfamiliar locations. Age groups also play a role in route preferences, with individuals aged 17-26 showing a higher preference for Google Maps. Hence considering data and statistics there is a higher probability of estimating that most of (70%) plus micromobility users in Sri Lanka would desire to use mobile application for their route navigation purposes [7].

D. Advantages and Challenges of Support Services Mobile Application

Development of a comprehensive support services mobile application for Micromobility users will be supportive for convenient and safer journeys. It helps to provide real time details of locations for micromobility vehicle parking areas to ensure the safety of the vehicle by avoiding vehicle theft. And the application will provide details of the locations of the charging stations for micromobility users who are using electricity powered micromobility vehicles, so they can travel more distance effortlessly by pedaling. This application will provide detailed locations of the bicycle and micromobility vehicle repair centers. Ensuring safer ride through the identification of bicycle tracks and cycle lanes which improve the bike ability (ability to ride) index [8]. This application will show the locations of the nearest bicycle lane and tracks, therefore micromobility users can ride safely. Substituting car travel with active modes offers cities the prospect of becoming cleaner, healthier, and more pleasant places.

However, this positive shift also brings challenges. In Denmark and the Netherlands, where a substantial influx of bicycles has occurred, issues of congestion in bike lanes and insufficient bike storage capacity have arisen. The consequence has been the unauthorized parking of bikes in various locations, obstructing access to stores, blocking streets, and diminishing the attractiveness of shopping areas and public squares. Recognizing the need for convenient, secure, and covered bike parking, cities in Denmark, Denmark, and Germany are actively working to enhance the security of such facilities. The unplanned parking of bikes in public spaces not only poses obstacles for pedestrians on sidewalks but is also deemed by some as an unpleasant visual disruption. Therefore, the expansion of bike parking facilities is driven not only by the desire to cater to cyclists' needs but also to address the issue of disorganized bike parking. This situation mirrors the challenges seen with car parking in the USA, where the demand for parking spaces always seems to exceed the available supply. The focus on managing bike parking is underscored as crucial for maintaining the livability of city centers. This involves making concerted efforts to ensure that the growth in active transportation modes aligns with improved conditions on city streets. As cities strive to strike this balance, the importance of bike parking management becomes increasingly evident, contributing to the overall sustainability and appeal of urban centers. Cities aim to maintain accessibility to their city centers while capitalizing on the advantages of slower transportation modes, seeking to enhance overall street conditions. Among the crucial components contributing to these objectives, bike parking stands out. However, discussions on bike parking and associated challenges in areas such as stations, shopping malls, and city centers are conspicuously scarce [9]. But when comparing Sri Lankan context, bicycle parking areas limited to metropolitan areas. Therefore, the bike ability factors are limited.

The presence of bicycle infrastructures plays a crucial role in shaping the motivation and overall experience of cycling. Therefore, to promote bicycle usage, it is essential for bicycle infrastructure to be both appealing and convenient for riders. Such infrastructure can boost cycling rates by increasing visibility, improving convenience, enhancing safety measures, and minimizing points of conflict with automobiles [10]. Cycle lanes are one of the main bicycle infrastructures which are more important for micromobility users to make sure of a safety journey while reducing the interaction with other vehicles. However, all riders underscored certain factors that diminish their motivation for cycling. The predominant concern raised was the obstruction of bicycle lanes by other vehicles, with 86% of riders expressing that vehicles parked in bicycle lanes discourage them from cycling. These observations underscore the need for well-designed features in bicycle lane infrastructure. Additionally, 50% of respondents pointed out that aside from cyclists, automobile drivers also use these lanes for driving, creating unsafe conditions for cycling [8].

III. DESIGN OVERVIEW

A. Diagrams

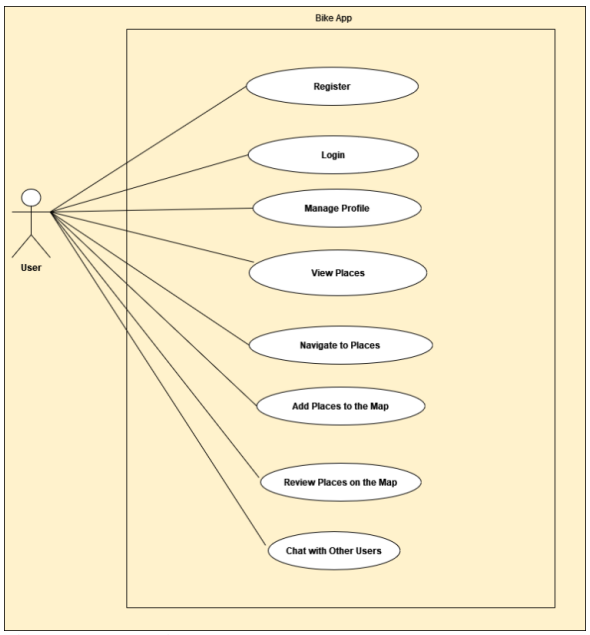


Fig. 1. Use Case Diagram

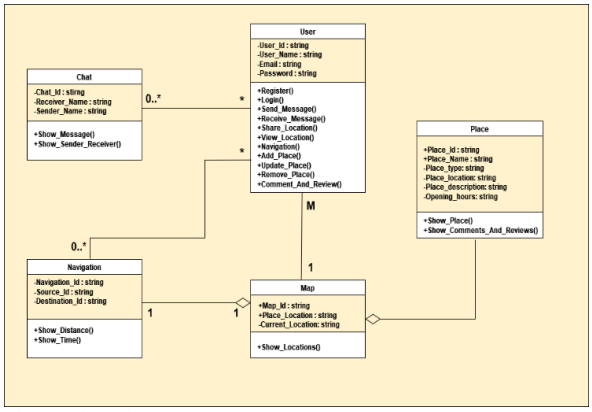


Fig. 2. Class Diagram

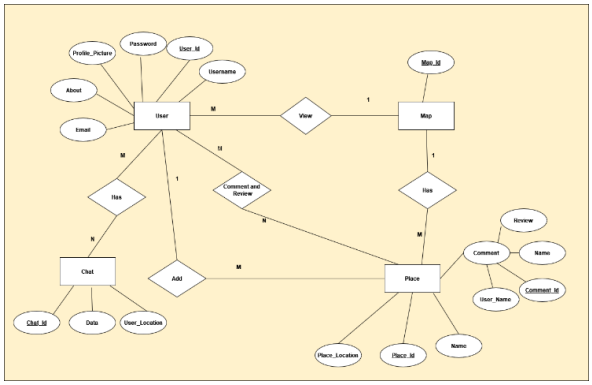


Fig. 3. Entity Relationship Diagram

Above entity relationship, use case and class diagrams describe the main functionalities of the proposed mobile application. The functional requirements of the mobile application consist with user login and registration, user able to manage his profile, displaying the user's current location, and locations of the other support services such as repair shops, charging stations, bicycle lanes and parking areas using the map feature of the application, enabling navigation function for user with the aid of Google Maps navigation. Additionally, users can contribute to the application by adding missing places to the map and those added places will be reviewed by other users to verify contributed place by commenting and voting. Finally enabling the messaging function with other users while being able to share the current location with their desired other users.

This proposed mobile application is developed using Flutter framework with dart language for development of mobile widgets. Firebase Authentication is used for user authentication a using Google signing and email password signing and 'Cloud Firestore', a NoSQL cloud database that allows for real-time data synchronization across users, which is a flexible, scalable database for mobile development. Firestore integrates seamlessly with other Firebase services like Firebase Authentication. 'Stream Builder' widget which is a powerful widget in Flutter framework that helps to manage data streams which can particularly use for building chat function of the mobile application. In here a stream is referred to as sequence of asynchronous events, for example, data from a chat server or a real-time database like Firestore can be represented as a stream. 'Share_pluse' Flutter plugin is used for the location sharing feature among users of the application. This plugin provides a simple way to share content, such as text, URLs, images, and files, with other apps installed on the user's device and useful for implementing features like sharing the user's current location via messaging apps, email, social media, etc.

In terms of Map functionality, Google Maps Platform is used because it offers several benefits and functionalities that are essential for the development of the features of the proposed mobile application. To obtain the current location of the user, the 'geolocator' plugin is used as it can be conveniently used with the flutter application, and which can obtain the current location of the user more accurately. To obtain the details of the support services such as bicycle parking areas, repair shops, bicycle lanes and charging stations, the 'Places API' which is provided by Google Maps Platform can be used to access information from the same database utilized by Google Maps. The information related to places which are not included in Google Places API will be gathered via crowd-sourcing feature which enables users to add places to the map and review those added place to confirm the verification about added places and also this feature will fulfilled by cloud Firestore database [11].

IV. DEVELOPMENT

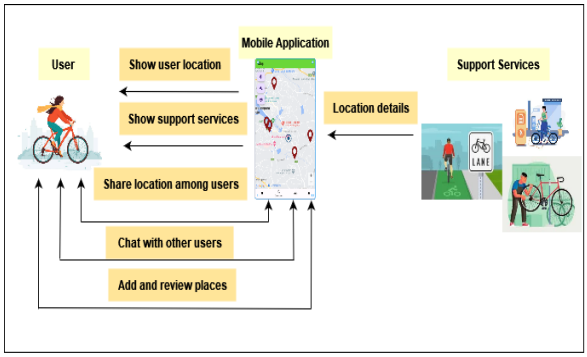
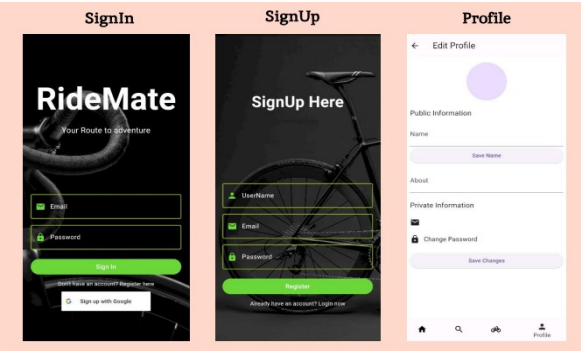


Fig. 4. Process Diagram of the system

The above diagram shows the process flow of the system. When the user successfully registered and logged to the system, user will land on the map screen in the application, in the map screen user able view his live location and the nearest locations support services which are including bicycle repair shops, bicycle parking areas, bicycle lanes and tracks, and charging stations for electrical micro-mobility users. This helps the user to figure out his current location and the nearest locations of required support services. If a user wants to navigate to a support service location or another desired location, this app facilitates this feature with the aid of the Google Maps application. When a location is selected and the navigate button is clicked, the application automatically starts Google Maps and initiates navigation. Users can add places to the map, comment and review on the added place. The bottom-line navigation bar facilitates the user navigating through the screens. The profile screen allows users to manage their information and view their chats with other users. Within the chat, users can share their current location with other users. Below are the figures of the proposed mobile application.



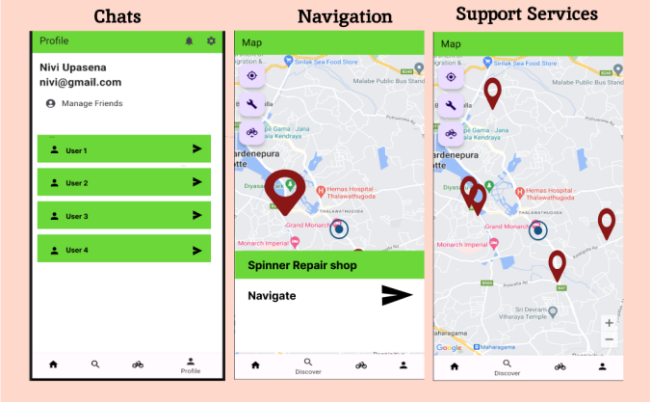


Fig. 5. Interfaces of the proposed application

V. CONCLUSION

This research paper highlights the development of a support services mobile application specifically designed for micromobility users in Sri Lanka. With the rapid urbanization of the country and the increasing adoption of micromobility solutions, this application addresses critical gaps in support services for users. The literature review revealed the significant benefits of micromobility, including its economic, environmental, and health advantages. However, it also identified the common challenges and pain points faced by micromobility users, such as the lack of centralized information on support services like secure parking, charging stations, and repair shops. The study highlights the benefits of the proposed mobile application of centralized support system integrates multiple functionalities to provide seamless user experience.

Once the proposed mobile application is developed, it undergoes testing phases to iteratively improve the functionalities of the mobile application by delivering the prototypes to users and gathering their feedback. This user-centered approach ensures that the final product effectively meets the functional and user requirements outlined in the research. Ultimately, the integration of support services into a mobile application represents a pivotal step forward in the micromobility landscape. By highlighting user safety, convenience, and satisfaction, this proposed solution has the potential to transform how individuals interact with micromobility options, making these services more accessible and reliable for all.

REFERENCES

[1] M. D. P. I. Jayasena and G. N. Samarasekara, 'Intention to Adopt Cycling using the Modified Theory of Planned Behaviour', Engineer: Journal of the Institution of Engineers, Sri Lanka, vol. 57, no. 2, pp. 89–98, May 2024, doi: 10.4038/engineer.v57i2.7615.

[2] S. Suntharalingam and L. Gunawardena, 'Cluster Analysis on Shared Micromobility Enthusiasts in Sri Lanka'.

[3] L. M. Braun et al., 'Short-term planning and policy interventions to promote cycling in urban centers: Findings from a commute mode choice analysis in Barcelona, Spain', Transp Res Part A Policy Pract, vol. 89, pp. 164–183, Jul. 2016, doi: 10.1016/J.TRA.2016.05.007.

[4] Maduranga, M. W. P., D. Ganepola, and R. P. S. Kathriarachchi. "Comparison of Trilateration and Supervised Learning Techniques for BLE Based Indoor Localization." ,2021.

[5] E. Pisoni, P. Christidis, and E. Navajas Cawood, 'Active mobility versus motorized transport? User choices and benefits for the society', Science of the Total Environment, vol. 806, Feb. 2022, doi: 10.1016/j.scitotenv.2021.150627.

[6] A. Francke, 'Cycling during and after the COVID-19 pandemic', Advances in Transport Policy and Planning, vol. 10, p. 265, Jan. 2022, doi: 10.1016/BS.ATPP.2022.04.011.

[7] 'What are the defining traits of the modern amateur cyclist? | Cycling Weekly'. Accessed: Feb. 09, 2024. [Online]. Available: https://www.cyclingweekly.com/fitness/what-are-the-defining-traits-of-the-modern-amateur-cyclist

[8] V. Garusinghe and A. Jayasinghe, 'NAVIGATION APP AND FUTURE SPATIAL THINKING: IMPACTS OF MOBILE NAVIGATION APP ON TRAVELER ROUTE CHOICE', 2021.

[9] N. Kankanamge, 'Examination of newly established bicycle lanes in Sri Lanka with special reference to Piliyandala and Katubedda', 2018.

[10] S. C. Van der Spek and N. Scheltema, 'The importance of bicycle parking management', Research in Transportation Business & Management, vol. 15, pp. 39–49, Jun. 2015, doi: 10.1016/J.RTBM.2015.03.001.

[11] C. Silva, J. Teixeira, and A. Proença, 'Revealing the Cycling Potential of Starter Cycling Cities', in Transportation Research Procedia, Elsevier B.V., 2019, pp. 637–654. doi: 10.1016/j.trpro.2019.09.113.

Protecting Vision in the Digital Age: Developing an IoT Solution for Monitoring Blue Light Intensity and Enhancing Eye Health

Thisura Samarakoon
Dept. of Computer and Data Science
NSBM Green University Town
Homagama, Sri Lanka
thisura.samarakoon@gmail.com

Kushani Kaushalya
Dept. of Computer and Data Science
NSBM Green University Town
Homagama, Sri Lanka
kushanikaushalya456@gmail.com

Srimal Fernando
Dept. of Computer and Data Science
NSBM Green University Town
Homagama, Sri Lanka
srimal.f@nsbm.ac.lk

Heshan Weerasinghe
Dept. of Information and System Sciences
NSBM Green University Town
Homagama, Sri Lanka
hkalhara777@gmail.com

Samadhee Samarasinghe
Dept. of Computer and Data Science
NSBM Green University Town
Homagama, Sri Lanka
ssadeesha11@gmail.com

Gowreeshan Karuneswaren
Dept. of Computer and Data Science
NSBM Green University Town
Homagama, Sri Lanka
kgowreeshan@students.nsbm.ac.lk

Tharindu Lakshan
Dept. of Network and Security
NSBM Green University Town
Homagama, Sri Lanka
tlakshan470@gmail.com

Chamal Rashmika
Dept. of Computer and Data Science
NSBM Green University Town
Homagama, Sri Lanka
lkrashmika@gmail.com

Isuru Sri Bandara Nishshanka
Faculty of Computing
NSBM Green University Town
Homagama, Sri Lanka
isuru.s@nsbm.ac.lk

***Abstract*— This research project aims to develop an Internet of Things (IoT) device designed to measure light intensity and color in indoor environments from reflective surfaces to artificial light sources and provide warnings when these levels are harmful to human eyes. The device will use sensors to detect changes in lighting conditions and communicate this data to a central system, which will analyze it and generate real-time alerts for users. This proactive approach helps reduce the risks associated with poor lighting conditions, thereby promoting better eye health and overall well-being.**

Keywords*— *Internet of Things (IoT), Light Intensity Measurement, Eye Health, Indoor Lighting Conditions

I. INTRODUCTION

The importance of proper lighting in indoor environments cannot be overstated. Especially considering the amount of time individuals spend indoors exposing from reflective surfaces to artificial light sources whether at home, work or school. Improper lighting characterized by excessive or insufficient light intensity and improper color temperatures can cause eye strain, headaches and long-term vision problems.

This research project seeks to address these issues by developing an IoT-based solution that continuously monitors lighting conditions in each space. By deploying a network of light sensors, the device will collect data on light intensity and color temperature, which will be processed to determine whether current lighting conditions are within safe and comfortable limits. When adverse conditions are detected, the system alerts users, allowing them to take corrective action, such as adjusting lighting or relocating themselves in the room.

The proposed solution will not only improve the comfort and productivity of users in indoor settings but will also contribute to long-term eye health. The project involves designing the hardware and software components of the IoT device, integrating them into an integrated system, and testing its effectiveness in real-world environments.

The primary objectives of this project are to,

1. Promote eye health and well-being.

2. Develop an accurate monitoring system.

3. Ensure real-time data processing and alerts.

4. User-friendly interface and notifications.

5. Provide cost-effective and scalable solution for the device.

II. RELATED WORKS

Widespread use of digital devices has led to a significant increase in screen time, raising concerns about possible adverse effects on eye health. Among the various factors that contribute to eye strain, blue light from screens and indoor lights have been identified as a significant problem. This literature review aims to explore the impact of blue light on eye health, the phenomenon of eye strain, and potential IoT-based solutions for monitoring and managing blue light exposure.

A. Impact of blue light on eye health

Blue light is a high energy visible light (HEV) with wavelengths between 400 and 495 nanometers. It has gained attention due to its potential to cause harm to human eyes. Research indicates that long exposure to blue light can lead to retinal damage and increase the risk of age-related macular degeneration (AMD) [1]. Also, blue light exposure is known to disrupt the 24-hour rhythm, affecting sleep quality and overall well- being [2].

Studies have shown that blue light can penetrate deep into the eye. This potentially causes oxidative stress on retinal cells [3]. This oxidative stress is linked to the generation of reactive oxygen species (ROS), which can damage cellular components and lead to visual impairment over time [4]. Understanding the biological mechanisms behind blue light-induced damage is crucial for developing effective preventive measures.

Researchers found that LED and cool white, fluorescent lighting significantly improved alertness and typing performance compared to warm white, fluorescent lighting [11]. This indicates that the type of light and its color temperature play a significant role in how comfortable and productive we feel in an indoor environment.

B. Digital eye strain and computer vision syndrome

Digital eye strain is also known as computer vision syndrome. It encompasses a range of symptoms including eye fatigue, dryness, blurred vision, and headaches [5]. The increasing reliance on digital devices for work, education, and entertainment has intensified the frequency of digital eye strain, affecting millions worldwide. The American Optometric Association (AOA) estimates that around 58% of adults experience symptoms of digital eye strain.

Factors contributing to digital eye strain include lengthy screen time, poor ergonomics, and imperfect lighting conditions [6]. Blue light exposure from screens is particularly concerning as it can worsen these symptoms, leading to discomfort and potential long-term vision problems. The need for comprehensive guidelines and interventions to mitigate eye strain has become more demanding in the digital age.

C. User awareness and behavior

Studies have shown that many users are not aware of the harmful effects of blue light and the importance of adopting protective measures [7]. Awareness campaigns and educational programs can significantly enhance users' understanding and encourage healthier screen and lighting habits.

Behavioral interventions, such as taking regular breaks using the 20-20-20 rule (every 20 minutes, look at something 20 feet away for at least 20 seconds), adjusting screen brightness, and using blue light filters can effectively reduce symptoms of eye strain [8]. Additionally, the adoption of ergonomically designed workstations and adherence to proper viewing distances can further reduce discomfort.

D. IoT solutions for eye health management

The introduction of the IoT offers promising solutions for real-time monitoring of blue light exposure. IoT- based devices equipped with sensors can measure blue light intensity emitted from digital screens and environmental sources. These devices can provide users with actionable feedback to adjust their screen settings and reduce exposure.

For instance, smart glasses embedded with blue light sensors can monitor real-time exposure and automatically adjust tint levels to filter out harmful blue light [9]. Similarly, IoT-enabled lighting systems can adjust ambient light to minimize blue light exposure during evening hours. Therefore, it is supporting better sleep hygiene [10]. The integration of IoT technology in managing blue light exposure presents a novel approach to improving eye health and mitigating eye strain.

E. Conclusion

All the literature highlights the significant impact of blue light on eye health and the prevalence of eye strain due to increased screen time. While awareness of these issues is growing, effective solutions for real-time monitoring and management of blue light exposure are still emerging. User awareness and behavior are critical in addressing these challenges, and IoT-based solutions show promise in bridging this gap.

III. CONCEPT OF THE SYSTEM

The Smart IoT-Based Light Quality Monitoring System for Enhancing Indoor Eye Health is designed to provide a comprehensive solution for monitoring and managing light conditions in indoor environments. The system integrates advanced sensor technology, microcontrollers, wireless communication modules, and user-friendly applications to ensure optimal lighting, thereby promoting better eye health and overall well-being. The system architecture is divided into three main modules: Module Sensing, Module Network Service, and Module Application, each playing a critical role in the overall functionality.

1. Module Sensing

The sensing module is responsible for capturing real-time data on light conditions. It consists of the following key components:

- Microcontroller (ESP8266): The ESP8266 microcontroller is used for processing sensor data and handling communication with the central system. It is chosen for its low power consumption, built-in Wi-Fi capabilities, and ease of integration with various sensors.
- Color Sensor (TCS34725): The color sensor detects the color temperature of the light. Different color temperatures can affect eye comfort and productivity, and this sensor helps identify any harmful variations in light quality.
- Buzzer and LED Indicators: These components provide immediate, on-site alerts to users when unfavorable light conditions are detected. The buzzer generates audible alerts, while the RGB LEDs offer visual indicators of the current light quality status.
- Power Supply: A 5V/3.3V power adapter or battery pack powers the microcontroller and sensors, ensuring consistent operation of the system.

2. Module Network Service

The network service module facilitates seamless communication between the sensing module and the central processing system. It includes:

- Wi-Fi Module (ESP8266): The ESP8266 Wi-Fi module enables wireless communication, allowing the microcontroller to send data to the central system over the internet. It supports the 802.11 b/g/n protocols, providing reliable connectivity for real-time data transmission.
- Cloud Storage and Processing Platform: Salesforce CRM platform is used as the cloud platform as it can be utilized for storing historical data and performing advanced analytics. This platform helps in long-term monitoring and trend analysis, providing deeper insights into lighting conditions and their impact on eye health.

3. Module Application

The application module provides a user-friendly interface for monitoring and managing the light quality data. It comprises:

- Web-Based or Mobile Application: This application displays real-time data on light intensity and color temperature, allowing users to monitor the current lighting conditions. It also provides alerts and notifications when unfavorable conditions are detected, enabling users to take corrective actions promptly.
- Data Visualization: The application includes features for visualizing the captured data, helping users understand the patterns and ranges of light conditions. It uses graphical representations to highlight any deviations from optimal lighting conditions, making it easier for users to adjust their environment accordingly.

The following visualizations demonstrate the functionality and effectiveness of the system:

A. Light Intensity and Color Temperature Over Time:

The graph below shows the variations in light intensity and color temperature over a specified period, illustrating the system's capability to monitor environmental light conditions continuously.

Fig. 1. Light intensity alerts over time graph

B. Alerts Visualization (Assuming alerts are triggered when light intensity > 900 lux)

The graph below highlights instances where the light intensity exceeded a predefined threshold, demonstrating the system's ability to identify and signal potentially harmful lighting conditions.

By integrating these modules into a cohesive system, the Smart IoT-Based Light Quality Monitoring System aims to provide a comprehensive solution for enhancing indoor eye health.

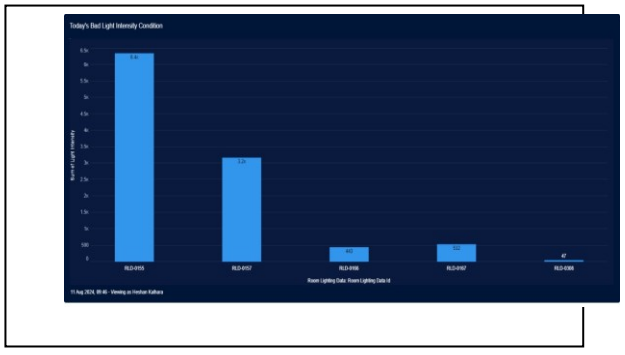


Fig. 2. Light intensity alerts over time graph

The proactive approach of real-time monitoring and alerts helps mitigate the risks associated with poor lighting conditions, promoting better eye health and overall well-being.

IV. SYSTEM MODEL

The system model integrates hardware, software, and cloud-based components to create a cohesive IoT solution for monitoring indoor lighting conditions. At the core of the system is the ESP8266 microcontroller, which interfaces with a TCS34725 RGB sensor to

measure light attributes and an OLED display to provide real-time data visualization. The hardware setup also includes a piezo buzzer that alerts users when the lighting deviates from the optimal range specified in user preferences. Data collected by the sensors is transmitted via Wi-Fi to the Salesforce cloud platform, where it is stored in custom objects designed to record both environmental data and user preferences. The system leverages REST API communication to ensure efficient data transfer and trigger notifications when necessary, thereby creating a dynamic feedback loop that helps users maintain a healthy visual environment. This model exemplifies the integration of IoT technologies with cloud services to enhance user experience and well-being.

Salesforce is a highly effective platform for health-related data collection systems, particularly when it comes to ensuring both operational efficiency and data security. Its cloud-based architecture enables scalable and real-time data management, making it ideal for handling large volumes of health information from various sources, such as IoT devices, mobile apps, and wearable technology. Salesforce excels in data security by adhering to stringent industry standards like HIPAA and GDPR, offering built-in encryption, two-factor authentication, and audit trails to safeguard sensitive patient information. The platform's granular access controls ensure that only authorized personnel can access specific datasets, significantly reducing the risk of unauthorized access or data breaches. Additionally, Salesforce provides a secure environment for integrating third-party applications and APIs, further enhancing its capability to collect, process, and analyze health data in a compliant and safe manner. With real-time analytics, customized dashboards, and automated workflows, Salesforce not only ensures data integrity and security but also empowers healthcare providers to make informed, timely decisions that improve patient outcomes while maintaining the highest standards of data privacy and security.

A. Technical Details

The IoT solution for monitoring blue light intensity and enhancing eye health involves a system built around the ESP8266 microcontroller, which serves as the core component for data processing and communication.

The system integrates a TCS34725 RGB sensor to measure RGB values, which are then used to calculate color temperature and blue light intensity. These measurements, along with the overall light intensity, are displayed on an OLED screen, providing real-time feedback to users. A small piezo buzzer acts as an alert mechanism, notifying users when the lighting conditions are outside the preferred parameters.

ESP8266's built-in Wi-Fi capability allows for seamless data transmission to the Salesforce cloud platform via REST API. This integration enables the storage and analysis of lighting data and facilitates the sending of email notifications to users when the lighting conditions are deemed poor, based on predefined user preferences.

B. Hardware Design

The hardware design centers on the ESP8266 microcontroller, which interfaces with various components to monitor and report lighting conditions.

The TCS34725 RGB sensor is connected to the ESP8266 using the I2C communication protocol, providing data on the light's color spectrum, from which blue light levels and color temperature are derived.

The OLED display, also connected via I2C, presents these values in real-time to users, ensuring they are constantly informed about the environmental lighting.

The system includes a piezo buzzer connected to a digital output pin on the ESP8266, which activates to alert users when the lighting does not meet the user-defined standards for optimal eye health.

The entire system is powered by a 5V regulated supply, ensuring stable operation of all components.

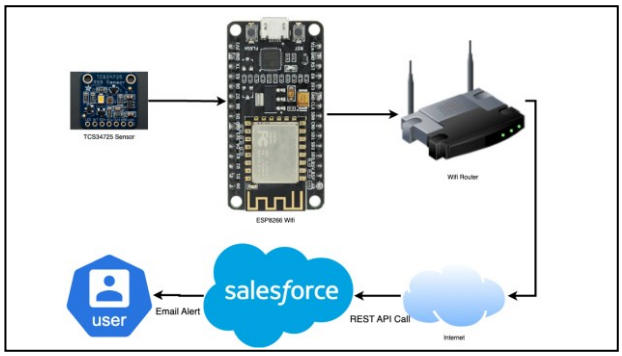


Fig. 3. Hardware design diagram

C. Database Design

The database design for this IoT solution involves the use of custom objects in the Salesforce cloud platform, specifically tailored to manage lighting data and user preferences.

The "Room Lighting Data" object captures real-time measurements from the sensors, including fields for Blue_Light_Level__c, Light_Intensity__c, and Color_Temperature_c. This object serves as the repository for all environmental data collected by the system.

In parallel, the "User Preferences" object stores individual user settings regarding acceptable lighting conditions, such as Preferred_Blue_Light_Max__c, Preferred_Color_Temperature_Min__c, Preferred_Color_Temperature_Max__c, and Preferred_Light_Intensity_Max__c. These preferences are used to evaluate the real-time data, triggering email notifications if the recorded lighting conditions deviate from the user's preferred settings.

This setup not only ensures the effective monitoring of indoor lighting conditions but also promotes user awareness and proactive management of their visual environment for better eye health.

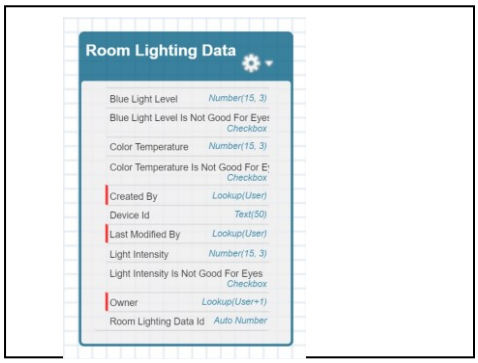


Fig. 4. Database tables in Salesforce

V. Materials and Methods

A. Microcontroller - ESP8266mod

The ESP8266mod Wi-Fi Module is a self-contained SOC with an integrated TCP/IP protocol stack, enabling microcontrollers to access Wi-Fi networks. It can host applications or offload Wi-Fi functions from other processors. Pre-programmed with AT command set code, it easily interfaces with Arduino devices. The cost-effective ESP8266mod features high-speed cache for enhanced performance and can connect via SPI/SDIO or I2C/UART interfaces, supported by a large community [12]. [13].

TABLE I. ESP8266 Specifications

Supported Standards	802.11 b/g/n, Wi-Fi Direct
Security	WPA/WPA2
Labor Input Volt. [V]	3.3/3 - 3.6
CPU	low power 32-bit @ 80 MHz
ROM-bootloader [KB]	64
RAM [KB]	64 (instructions) 96 (data)
Flash Memory [MB]	1
Serial Connection	SPI, I2C, UART
Size [nn]	25 x 38 x 5
Energy Consumpt. [mA]	~60 - 215 in exec. ~0.0009 in standby

B. Color Sensor – TCS34725

The TCS34725 sensor detects light with an 8x8 photodiode array, converting light energy into current, then into frequencies proportional to light intensity. It measures colour by analyzing light values with a microcontroller to calculate RGB values. This low-cost, high- sensitivity sensor is used in lighting monitoring systems for accurate, real-time spectral detection [14].

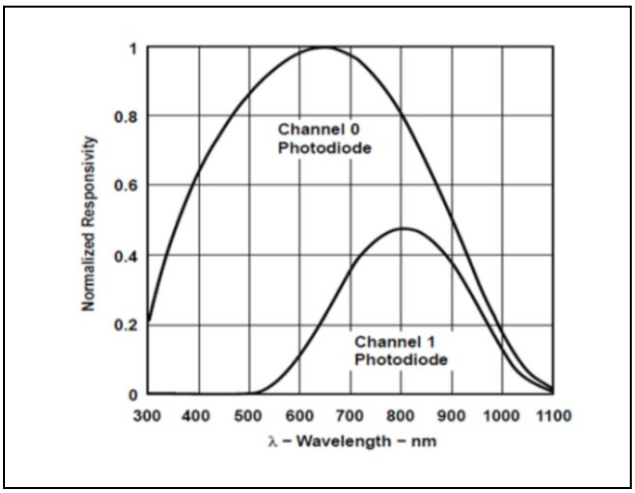


Fig. 5. Spectral responsivity of the TCS34725 sensor

C. Buzzer - Small piezo buzzer

In an eye protection IoT device, piezoelectric buzzers serve as acoustic signal generators. Their models, derived from data sheets and measurements, are used to simulate user circuits like CMOS-AMV, RC phase shifters, and oscillators with various sensors (force, phototransistor, NTC). By understanding the piezo element's characteristics and vibrational modes through systematic electrical modeling, the buzzer's frequency response and behavior in the required frequency range can be accurately predicted and incorporated into the device design [15] [16].

VI. Discussion and Results

The Smart IoT-Based Light Quality Monitoring System for Enhancing Indoor Eye Health aims to provide a holistic approach to maintaining optimal lighting conditions indoors. This system leverages advanced sensors, microcontrollers, and a user-friendly interface to continuously monitor light intensity and color temperature, thus ensuring an environment conducive to eye health.

A. Key Findings

1. Effectiveness of the Sensing Module

• The TCS34725 color sensor proved highly effective in capturing real-time lighting conditions. The precision and reliability of this sensor ensure accurate measurements, which are crucial for maintaining optimal lighting.

• Immediate on-site alerts through buzzers and OLED indicators were instrumental in informing users

of unfavorable lighting conditions, prompting timely corrective actions.

2. Network Service Reliability

• The ESP8266 Wi-Fi module provided robust and reliable connectivity, enabling seamless data transmission to the central processing system. The option to use cloud storage for historical data and advanced analytics added significant value by facilitating long-term monitoring and trend analysis.

• The integration of cloud services, although optional, offered deeper insights into lighting patterns and their impact on eye health, allowing for more informed decisions regarding environmental adjustments.

3. User Interface and Experience

• The web-based and mobile applications designed for this system offered an intuitive interface for real-time monitoring. Users could easily visualize light intensity and color temperature data, and receive alerts for any deviations from optimal conditions.

• Data visualization tools within the application, including graphs and charts, were highly effective in helping users understand lighting patterns and adjust their environment accordingly.

B. Performance Evaluation

The system was subjected to rigorous testing in various indoor settings to evaluate its performance on several key metrics.

• Response Time: The system demonstrated rapid detection and response capabilities, with minimal latency between sensor data capture and user notification. This quick response is critical in maintaining optimal lighting conditions and preventing eye strain.

• Detection Accuracy: The sensors' high accuracy in measuring light intensity and color temperature ensured reliable data, essential for effective light quality management. The accuracy of the TCS34725 color sensor in detecting harmful variations in light quality was particularly noteworthy.

• User Satisfaction: Feedback from users indicated a high level of satisfaction with the system's performance. The immediate alerts and detailed data visualizations were particularly appreciated, as they provided actionable insights and facilitated better management of indoor lighting.

C. Visualizations and Alerts

• Light Intensity and Color Temperature Over Time: Graphs depicting light intensity and color temperature variations over specified periods highlighted the system's continuous monitoring capability. These visualizations demonstrated how the system maintained lighting within optimal ranges, adapting to changes in the environment.

• Alerts Visualization: Instances where light intensity exceeded predefined thresholds were effectively highlighted, showcasing the system's ability to identify and signal potentially harmful lighting conditions. This functionality is crucial for preventing prolonged exposure to suboptimal lighting.

D. Challenges and Improvements

During the testing phase, several challenges were identified.

• Environmental Variability: Variations in environmental conditions, such as the presence of natural light or reflections, sometimes affected sensor readings. Future iterations of the system could incorporate advanced algorithms to compensate for such variability, enhancing accuracy.

• User Adaptation: Educating users on the system's importance and functionality was vital. Despite initial hesitations, user adaptation improved significantly with comprehensive instructional materials and awareness campaigns.

• Cost-effective Cloud Solutions: Salesforce in cost wise is very high for an implementation like this project, but considering the security, fast development time and flexibility Salesforce is a better solution. But always we can develop the system using open source resources also.

The Smart IoT-Based Light Quality Monitoring System offers a comprehensive solution for enhancing indoor eye health. Its proactive real-time monitoring and alert capabilities ensure optimal lighting conditions, mitigating the risks associated with poor lighting. Through continuous evaluation and user feedback, the system has proven effective in various indoor environments, promoting better eye health and overall well-being.

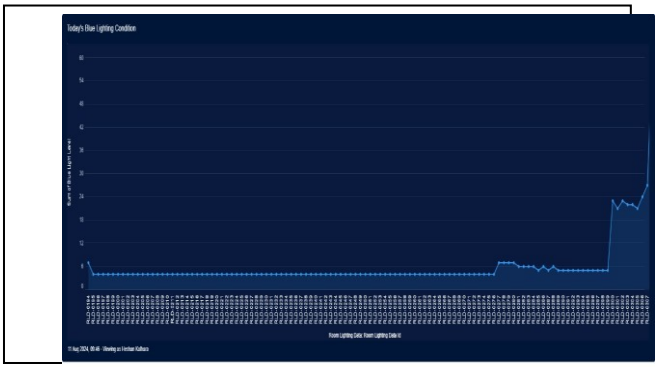


Fig. 6. Blue light condition for a day graph

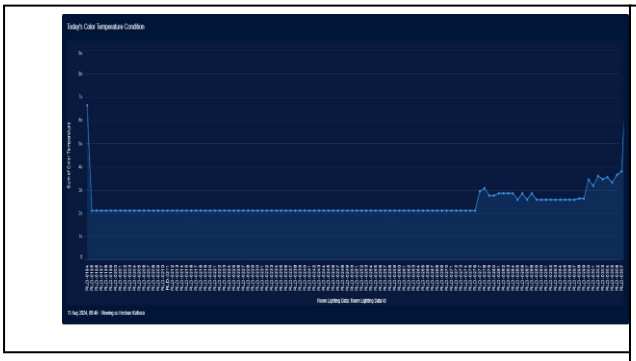


Fig. 7. Color temperature condition for a day graph

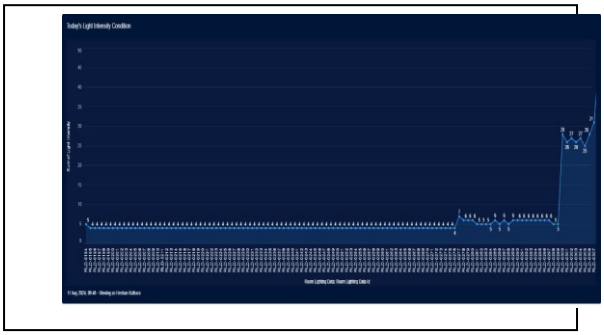


Fig. 8. Light intensity of a day graph

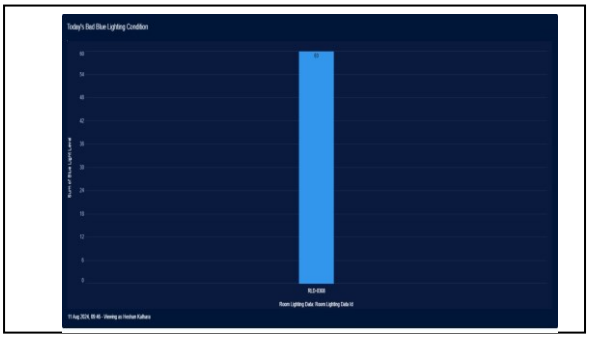


Fig. 9. Bad blue lighting condition for a given day graph

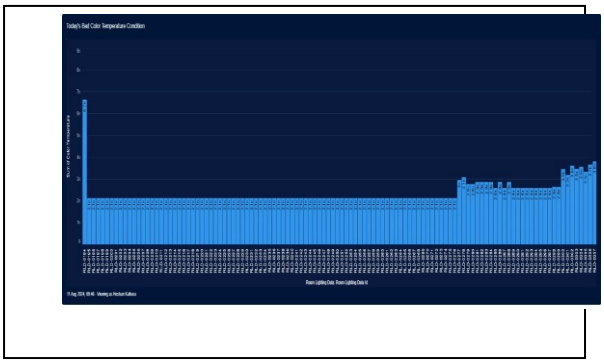


Fig. 10. Bad color temperature condition for a given day graph

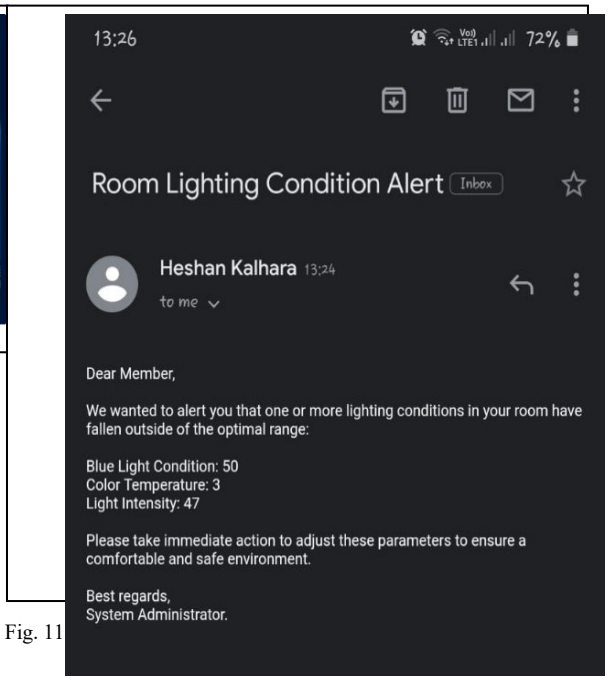


Fig. 11

VII. CONCLUSION

In conclusion, the developed IoT solution offers a comprehensive approach to monitoring indoor lighting conditions, focusing particularly on blue light intensity—a critical factor for eye health.

By leveraging the ESP8266 microcontroller, TCS34725 RGB sensor, OLED display, and a piezo buzzer, the system provides real-time feedback and alerts on lighting parameters, including blue light levels, color temperature, and overall light intensity.

The integration with the Salesforce cloud platform enables security, efficient data management and user-specific alerting mechanisms, ensuring that users are promptly informed of any deviations from their preferred lighting conditions.

This system not only enhances user awareness of their visual environment but also encourages healthier lighting practices. The implementation demonstrates a practical application of IoT in promoting eye health, with potential for further refinement and expansion, including integration with additional sensors and the development of more sophisticated data analytics capabilities. This work lays the groundwork for future advancements in smart lighting systems and their role in mitigating the adverse effects of modern indoor lighting on eye health.

REFERENCES

[1] F. Behar-Cohen, C. Martinsons, F. Viénot, G. Zissis, A. Barlier-Salsi, J. P. Cesarini, ... & D. Attia, "Light-emitting diodes (LED) for domestic lighting: Any risks for the eye?," Progress in Retinal and Eye Research, vol. 30, no. 4, pp. 239-257, 2011.

[2] M. Chang, D. Aeschbach, J. F. Duffy, & C. A. Czeisler, "Evening use of light-emitting eReaders negatively affects sleep, circadian timing, and next-morning

alertness," Proceedings of the National Academy of Sciences, vol. 112, no. 4, pp. 1232-1237, 2015.
[3] M. Boulton, M. Rózanowska, & B. Rózanowski, "Retinal photodamage," Journal of Photochemistry and Photobiology B: Biology, vol. 64, no. 2-3, pp. 144-161, 2001.
[4] Sliney, "How light reaches the eye and its components," International Journal of Toxicology, vol. 22, no. 6, pp. 501-509, 2016.
[5] M. Rosenfield, "Computer vision syndrome: a review of ocular causes and potential treatments," Ophthalmic and Physiological Optics, vol. 31, no. 5, pp. 502-515, 2011.
[6] J. E. Sheedy, J. R. Hayes, & J. Engle, "Is all asthenopia the same?," Optometry and Vision Science, vol. 80, no. 11, pp. 732-739, 2003.
[7] H. J. Lee, M. W. Cho, & S. H. Kim, "Awareness and behavior regarding blue light exposure among digital device users," Journal of the Korean Ophthalmological Society, vol. 58, no. 4, pp. 344-350, 2017.
[8] S. H. Yen, M. Y. Sun, & H. S. Lee, "Effectiveness of educational interventions in reducing digital eye strain among digital device users," Journal of Educational Technology & Society, vol. 21, no. 3, pp. 215- 226, 2018.
[9] X. Li, D. Wen, Q. Du, Y. Fu, H. Wu, & S. Chen, "Smart glasses for real-time blue light filtering to alleviate digital eye strain," IEEE Internet of Things Journal, vol. 5, no. 6, pp. 5102-5111, 2018.
[10] Wood, M. Rea, B. Plitnick, & M. Figueiro, "Light level and duration of exposure determine the impact of self-luminous tablets on melatonin suppression," Applied Ergonomics, vol. 44, no. 2, pp. 237- 240, 2013.
[11] Kazemi, R., Choobineh, A., Taheri, S., & Rastipishe, P. (2018). Comparing task performance, visual comfort and alertness under different lighting sources: An experimental study. EXCLI Journal, 17, 1018- 1029. doi:10.17179/excli2018-1676
[12] R. Shinde1, "Review on Real Time Approach for Smart Agriculture Based on IoT and," International Journal of Computer Sciences and Engineering , 2019.
[13] G. O. P. C. E. &. Q. G. Bujari, "Wearable sensor networks: A measurement study," 2021.
[14] S. Hussein, LOW COST PORTABLE URINE ANALYSIS SYSTEM DESIGN BASED ON COLOR, International Marmara Sciences Congress, 2020.
[15] P. Baumann, Piezoelectric Buzzer, 2023.
[16] D. a. B. P. a. G. B. a. G. R. Banerjee, "Deriving an electrical model of a two terminal piezoelectric buzzer from its impedance response," in Applied Signal Processing Conference (ASPCON), 2018.

Cybersecurity Awareness and Challenges in Sri Lankan Small and Medium-Sized Enterprises: A Comprehensive Survey

Tharun Alvitigala
Department of Computer Security and Network System
National School of Business Management
Colombo, Sri Lanka
tharunalvitigala@gmail.com

Madusanka Mithrananda
Department of Computer Security and Network Systems
National School of Business Management
Colombo, Sri Lanka
madusanka.m@nsbm.ac.lk

***Abstract*— Small and medium-sized enterprises (SMEs) are integral to both global and Sri Lankan economies. Despite their significant presence, cybersecurity research rarely focuses on SMEs, leaving them vulnerable to cyber threats. Economic limitations and crises have further degraded the security of SMEs, occasionally exposing them to cyberattacks. This study aims to identify the level of cybersecurity awareness among Sri Lankan SMEs, the precautionary measures they have implemented, and the types of cyberattacks they have experienced. Conducted against a backdrop of limited research on this topic in Sri Lanka, this survey provides a comprehensive understanding of the current cybersecurity environment within SMEs. By highlighting the gaps in security practices and awareness, the research seeks to inform better protective measures tailored to the needs of these businesses. The findings underscore the critical need for enhanced cybersecurity protocols and targeted awareness programs to safeguard SMEs from escalating cyber threats. This study fills a crucial gap in existing literature and offers valuable insights into the cybersecurity challenges and responses of SMEs in Sri Lanka.**

Keywords*— *Small-to-medium business, Cyber threats, Cybersecurity awareness, Precautionary measures, Cybersecurity challenges

I. INTRODUCTION

Small and medium-sized organizations (SMEs) constitute a vital segment of the global economy, contributing significantly to employment and innovation. They play a major role in the Sri Lankan economy, constituting 75% of all active businesses. These companies provide 45% of employment while contributing 52% of the country's gross domestic product (GDP)[1].

With the rise of complex cybersecurity threats in the modern digital age, it is more important than ever for SMEs to implement a network security solution. SMEs typically face distinct obstacles compared to larger organizations, primarily due to limited resources and specialized knowledge. This security gap makes them attractive targets for cybercriminals. Common issues SMEs encounter include inadequate disaster recovery or business continuity strategies, limited awareness of the latest cyber-attacks, constrained IT capabilities, and insufficient cybersecurity investments[2].

Sri Lanka's position in global cybersecurity rankings highlights the challenges faced by SMEs. In 2020, Sri Lanka ranked 83rd out of 194 countries assessed by the International Telecommunications Union (ITU), securing the 15th position within the Asia Pacific region. Despite progress, cybersecurity among SMEs remains understudied due to several key factors. First, national policies often prioritize larger enterprises and government institutions, leaving SMEs neglected due to their smaller scale and perceived lower risk. Secondly, SMEs in Sri Lanka typically face financial and technical resource constraints, viewing cybersecurity as a cost rather than an investment, and prioritizing core business operations over IT security. Underreporting of cyber incidents further exacerbates the issue, as many SMEs fail to recognize attacks or prefer to remain silent due to reputational concerns, leading to a lack of data for researchers. Moreover, there is a notable gap in specialized research focused on SME cybersecurity, with most research efforts concentrating on larger entities or global trends. This lack of data hinders the development of proper solutions to safeguard SMEs against cyber threats, leaving them vulnerable to potential attacks[3]–[5].

Finally, the lack of internal cybersecurity expertise within SMEs creates vulnerabilities and hampers the development of both practical defenses and academic insights, leaving SMEs significantly underrepresented in cybersecurity discourse. Addressing these issues is crucial for enhancing the cybersecurity resilience of Sri Lankan SMEs[4].

This survey focuses on various types of SMEs across different industries in Sri Lanka, examining how they encounter and address different cybersecurity threats. It investigates the solutions these enterprises employ to minimize or prevent cyberattacks, the specific types of cyber threats they have faced, and the level of cybersecurity awareness among employees along with the ―budget allocations for cybersecurity. Thereby, this the survey aims to provide a comprehensive understanding of the cybersecurity landscape among SMEs in Sri Lanka, highlighting the challenges and practices within this critical sector.

II. LITERATURE REVIEW

An overview of existing literature on cybersecurity awareness highlights its critical importance for SMEs. Effective cybersecurity measures are essential to protect sensitive data, maintain customer trust, comply with regulations, prevent disruptions, mitigate financial losses, safeguard intellectual property, and enhance overall business resilience and competitiveness[6]. Although large-scale companies have a larger attack surface, they have dedicated security operations center (SOC) and network operations center (NOC) teams to control situations. In contrast, less experienced hackers often target SMEs to hone their skills.

A. Cyberattacks experienced by SMEs: Global

The rapid advancement of technology in recent years has led to a significant increase in the frequency of cyberattacks, with approximately 70% of internet-connected devices being vulnerable to such threats[7],[8]. In 2020, small businesses experienced over 700,000 cyberattacks, resulting in damages amounting to 2.8 billion dollars. These figures continue to increase each year[9].

Figure 1 shows the type of attacks experienced by SMEs in globally. (Figure 1) The majority (70%) of global cyber-attacks are perpetrated by external parties, with 16% being done by internal actors. Of these attackers, 86% are financially motivated, while the remaining 14% have other motivations, including ideology, state sponsorship, grudges, or simply for pleasure[10].

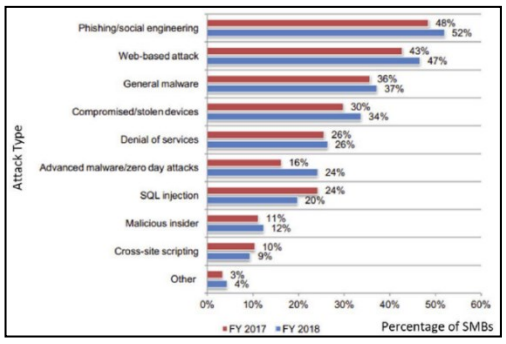


Fig. 1. Types of attacks experienced by SMBs [10]

Apart from Figure 1 listed attacks, open Wi-Fi highlights a significant danger and underscores the importance of employing security measures. A large-scale study conducted in 15 major airports worldwide found that two-thirds of users' private information had been leaked through public Wi-Fi in airports. Additional studies conducted in London and China found that sensitive information including income, investment details, usernames, and passwords was leaked during the use of public Wi-Fi[11].

The shift to remote work during the COVID-19 pandemic has led to a significant rise in cybercrime, leaving small businesses increasingly vulnerable to ongoing attacks[12], [13].

According to Small Business Administration (SBA) survey, the cost of a data breach can exceed up to $200,000 per incident. Additionally, 20% of customers have ended their cooperation with affected SMEs. Bartik's findings indicated that 60% of SMEs close permanently within six months following a cyberattack[14]. Moreover, 71% of companies with fewer than 100 employees are considered attractive targets for cybercriminals[9].

B. Cyberattacks experienced by SMEs: Sri Lanka

A joint study conducted by the United States Agency for International Development (USAID) and the Ceylon Chamber of Commerce (CCC) has found that a significant 88% of SMEs engaged in export activities in Sri Lanka have fallen victim to cyberattacks[15]. Another study in Sri Lanka, revealed that nearly 80% of SMEs had experienced cyberattacks or data breaches, with 72% reporting business disruptions as a result[16].

III. METHODOLOGY

This survey, involving 59 SMEs, was conducted to assess the cybersecurity threat landscape in Sri Lanka, with a focus on the experiences of SMEs. The primary objectives were to identify the types of cyberattacks affecting Sri Lankan SMEs, determine their monthly expenditure on cybersecurity, and evaluate the features included in their security budgets.

This study employed a quantitative research design. The overall strategy was descriptive, aiming to gather detailed information about the cybersecurity experiences of SMEs in Sri Lanka.

SMEs were categorized solely based on the number of employees due to confidentiality concerns. Small-scale companies were defined as having 1-30 employees, while medium-scale companies were defined as having 31-150 employees. This classification was necessary because SMEs' asset values and annual revenues were not disclosed for security reasons. Companies were selected randomly to ensure fairness and minimize bias, and all selected companies used the internet for daily operations, including financial transactions.

Data were collected through a questionnaire consisting of 24 questions designed to gain a comprehensive understanding of the companies. Two methods were employed for data collection: first, by visiting the companies and conducting the survey in person; and second, by sending the questionnaire via a Google Forms link for respondents to complete online. Data quality assurance was maintained by pre-testing the questionnaire and ensuring consistent data collection procedures.

The collected data were analyzed using statistical methods to identify patterns and trends. Descriptive statistics were employed to summarize the data. Data analysis was conducted using Microsoft excel software.

IV. RESULTS

A. Firm Demographics

The survey encompassed companies of varying sizes, with the number of employees ranging from 1 to 150. This provides a comprehensive understanding of how cybersecurity threats and measures impact organizations with different workforce scales. Majority of the respondent organizations had more than 30 employees. (Figure 2)

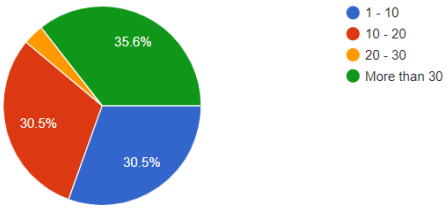
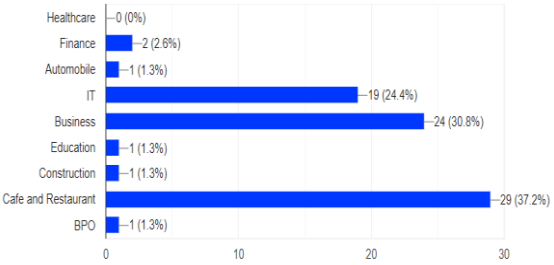


Fig. 2. Number of employees

The survey collected responses across various industry sectors. The cafe and restaurant sector had the most significant engagement, with 29 responses, constituting 37.2% of the total. (Figure 3). The business sector had 24 responses, making up 30.8% of the total. IT sector provided 19 responses, representing 24.4% of the total. The finance sector contributed two responses, accounting 2.6% of the total. Education, Construction, and BPO sectors each contributed one response, accounting for 1.3% each.

Fig. 3. Type of industry

Determining the organizational age helps identify the



time needed for companies to stabilize and adopt cybersecurity practices. Most companies (55.9%) are in the development stage (1-5 years), 15.37% are in the survival and growth period (5-10 years), 11% are in maturity, and 11.9% are in a stable environment (Figure 4).

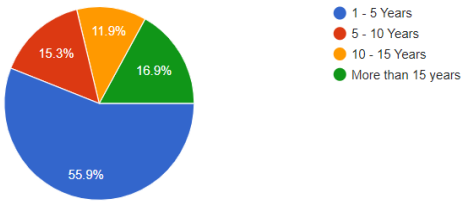


Fig.4. The number of years that the company is on business

B. Evaluating security solutions and budget allocation in organizations

The survey results revealed that only 44.1% (26) of respondents had security solutions implemented in their company. Among these 26 organizations, the following measures were reported: 57.7% (15) had physical backups, 69.2% (18) had server backups, 73.1% (19) used cloud backups, 96.2% (25) employed firewalls, 65.4% (17) had antivirus software, 38.5% (10) used access control, and 46.2%

(12) had intrusion prevention/ detection system (IPS/IDS) in place. (Figure 5)

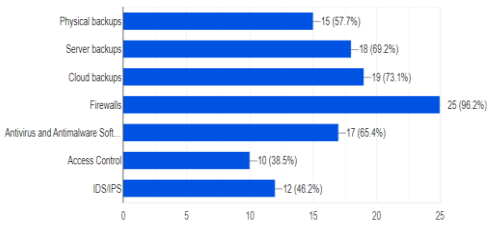


Fig.5. Types of security solutions

Of the 26 SMEs with security solutions, the majority (38.5%; 10 respondents) allocated a budget exceeding LKR 200,000/=. Other budget allocations were as follows: 3.8% (1 respondent) had a budget of less than LKR 50,000/=; 11.5% (3 respondents) allocated between LKR 50,000/= and 100,000/=; 23.1% (6 respondents) had a budget between LKR 100,000/= and 150,000/=; and 23.1% (6 respondents) allocated between LKR 150,000/= and 200,000/= (Figure 6).

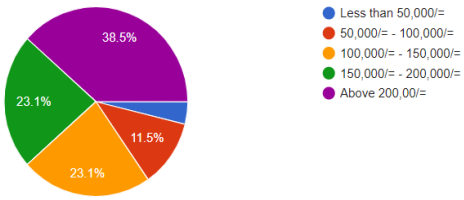


Fig. 6. Financial resources allocated

C. Preferences for affordable security solutions and desired features by all respondents

Regardless of whether they currently have a security solution or not, 92% of respondents preferred a more affordable option with enhanced security features.

The majority of respondents preferred the implementation of firewall solutions (65.9%), followed by IPS systems (63.4%) and anomaly detection systems (48.8%). IDS and rogue access point detectors were equally preferred by 40.2% of respondents. Virtual private networks (VPNs) and bandwidth monitoring systems were the least favored options. (Figure 7)

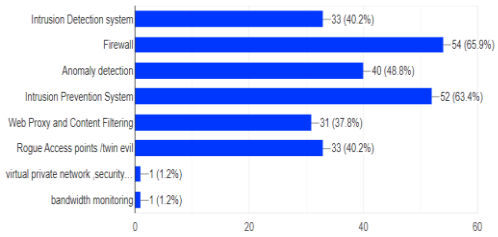


Fig. 7. Preferred security features among respondents

In contrast, the majority of respondents (63.2%; 36 respondents) preferred an affordable security solution costing a budget between LKR 50,000/= to 100,000/=, Meanwhile, 24.6% (14 respondents) preferred a solution less than LKR 50,000/=. and only 12.3% (7 respondents) indicated a budget exceeding LKR 100,000/= (Figure 8).

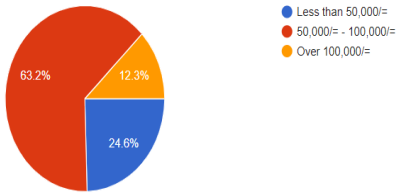


Fig. 8. Estimated budget for preferred solution

D. Cybersecurity attacks faced by respondents

Of the 59 respondents, 57.6% (34 respondents) have faced cybersecurity attacks.

The majority of companies reported experiencing phishing attacks, which accounted for approximately 77.4% of incidents. Following this, 51.6% of companies experienced malware and ransomware attacks. DoS attacks were reported by 38.7% of respondents, while insider threats were encountered by 16.1% of companies. Additionally, SQL injections, credential surfing, rogue hotspots and Wi-Fi, scam calls, bandwidth issues, and other cyber scams were each reported by 22.4% of respondents (Table 1).

TABLE 1. TYPE OF CYBERSECURITY THREATS FACED

Types of Cybersecurity threats	Incidents reported	
	Count	*Percentage*
Denial of service	12	38.7%
Phishing Attacks	24	77.4%
Malware and Ransomware	16	51.6%
Inside threats	5	16.1%
Rouge access points	4	9.7%
SQL injections	1	3.2%

Types of Cybersecurity threats	Incidents reported	
	Count	*Percentage*
Credential stuffing	1	3.2%
Scam calls / cyber scams	2	6.4%
Market manipulations	1	3.2%
Bandwidth anomalies	1	3.2%
System exploits	1	3.2%

E. Availability of a dedicated a knowledgeable IT staff in the responded SMEs

Of the 44.1% (26 respondents) had a dedicated IT staff.

Cybersecurity awareness among IT staff was rated on a scale from 1 (low) to 5 (high). The majority of respondents (27.1%; 16 respondents) rated it as 2, while 22% (13 respondents) rated it as 1. Conversely, 20.3% (12 respondents) rated their staff's cybersecurity awareness as high (5) (Figure 9).

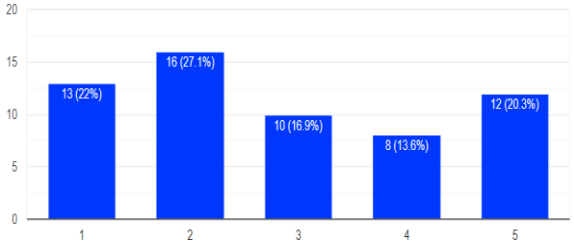


Fig. 9. Cybersecurity awareness among staff

V. DISCUSSION

A. Industry Engagement and Awareness

All the cafes and restaurants (29 responders) that participated in the survey offered free Wi-Fi hotspots for customers, enhancing their experience while on the premises. In contrast, companies in other sectors, such IT (19 responders) and business (5 responders), had multiple departments sharing the same Wi-Fi without any access points or access controls. While most IT and business companies demonstrated a solid understanding of cybersecurity and its importance, some lacked dedicated staff for this purpose. The survey revealed that education, automobile, construction companies, cafes and restaurants exhibited a significant lack of cybersecurity awareness. Conversely, the BPO and finance sectors showed a thorough understanding of cybersecurity.

Companies with 1-10 employees (30.5%) exhibited less dedication to having specialized IT staff for cybersecurity protection, whereas companies with more than 30 employees tend to have dedicated IT staff to safeguard their internal systems.

Most companies were in the development stage (1-5 years), aligning with the finding that over half (55.9%) lacked any cybersecurity solutions, suggesting a potential resource constraint during it's crucial growth period. Companies with solutions prioritized cloud backups (96.2%) and firewalls (65.4%), indicating an understanding of basic data protection needs. Antivirus adoption (38.5%) and access control (46.2%) were relatively lower, suggesting potential gaps in endpoint and user access security.

The budget distribution revealed a clear preference for affordable solutions. A significant portion (38.5%) allocated over LKR 200,000/=, indicating a willingness to invest in robust security when resources allowed.

B. Desired Security Features and Budget Constraints

The preference for affordable, tailored solutions aligns with the development stage (1-5 years) and budget limitations of most SMEs.
The desired features list (firewalls, IPS/IDS, anomaly detection, content filtering, VPN) reflected a comprehensive security posture, encompassing network protection, intrusion detection, data security, and access control. The proposed budget range (LKR 50,000 - 100,000/=) highlighted the need for cost-effective solutions that addressed these essential security needs.

C. Cybersecurity Attacks and IT Expertise

Over half (57%) of the companies had faced cyberattacks, including phishing (77%), malware/ransomware (51%), DoS attacks (38%), and cyber scams (22%). This underscored the prevalence of cyber threats targeting SMEs in Sri Lanka. This survey also identified a lack of dedicated IT staff (60%) and low cybersecurity awareness, further exacerbate the vulnerability of these companies.

A major issue identified was that most cafes and restaurants offered open Wi-Fi to customers while neglecting the protection of their internal systems, potentially exposing them to external threats and increasing the risk of rouge access points.

The security of free access points posed significant disadvantages[17]. Unprotected access points could allow hackers to access credit cards information, bank accounts, and other personal financial data[11], [18]. Public Wi-Fi networks, due to their broadcast nature, were inherently vulnerable to data

interception, making it possible for individuals nearby to access transmitted information. This could be exploited through various methods such as

war chalking, Wi-Fi mooching, joyriding, war driving, piggybacking, or hitchhiking[19], [20].

Sri Lankan SMEs face significant cybersecurity challenges due to limited awareness, gaps in national policies and resource constraints compared to countries with developed economies. They remain underserved by the government and banking sector in terms of financial support and targeted policies. The National Cybersecurity Strategy (2019–2023), primarily focus on larger enterprises and critical infrastructure, neglecting SMEs. Countries with robust cybersecurity frameworks often provide SMEs with tailored financial support, affordable cybersecurity tools, and mandatory reporting requirements, resulting in better protection and awareness. In contrast, Sri Lankan SMEs lag behind in implementing basic defenses and a lack of detailed research due to underreporting of cyber incidents leaves gaps in understanding the threat landscape. Addressing these gaps would greatly improve the cybersecurity resilience of Sri Lankan SMEs, enabling them to protect against the growing threats in today's digital landscape[4], [21], [22].

Although initiatives such as the Digital Economy Strategy promote digital transformation, they lack specific cybersecurity measures for SMEs. Government efforts, including ICTA and Sri Lanka CERT, focus on national cyber defense, while banks prioritize larger institutions, offering little assistance to SMEs. Without dedicated financial support and SME-specific cybersecurity programs, these businesses remain vulnerable to increasing cyber threats[4], [22]–[24].

VI. CONCLUSION

The study highlighted a concerning gap in cybersecurity awareness among Sri Lankan SMEs. SMEs can improve the security posture with following recommended actions without significant increase in their security budget. Improve employee cybersecurity awareness using regular training on current security threats to their respective business scope. Implementing basic cybersecurity measures such as firewalls, antivirus software, and secure Wi-Fi access points along with device hardening. Segment and password-protected the internal networks. Prioritizing investments in cloud backups and implementing opensource security applications such as open-source IDS, can help improve security for low cost. SMEs Could seek partnerships with local universities and IT professionals for expertise and guidance. Most importantly it is suggested to allocate a budget by doing a risk assessment with their business growth. These steps will improve data privacy and business continuity in SMEs.

ACKNOWLEDGMENT

The authors acknowledge the participating companies for providing their inputs for the survey.

REFERENCE

[1] S. Kulasinghe, L. Han, T. Hoshino, A. Rathi, and D. Lambert, "Catalyzing Small and Medium-sized Enterprise Venture Capital in Sri Lanka: ADB SOUTH ASIA Working Paper Series," 2018.

[2] W. S. Admass, Y. Y. Munaye, and A. A. Diro, "Cyber security: State of the art, challenges and future directions," *Cyber Secur. Appl.*, vol. 2, p. 100031, 2024, doi: 10.1016/j.csa.2023.100031.

[3] ITU, "Global Cybersecurity Index -2020: Measuring commitment to cybersecurity," Switzerland, 2021. Accessed: Oct. 12, 2024. [Online]. Available: https://www.itu.int/dms_pub/itu-d/opb/str/D-STR-GCI.01-2021-PDF-E.pdf.

[4] SLCERT, "National Cyber Security Strategy of Sri Lanka (2024 – 2027)," 2024. Accessed: Oct. 12, 2024. [Online]. Available: https://cert.gov.lk/wp-content/uploads/2024/03/NCSS-Draft-V14.pdf.

[5] DCS, " Computer Literacy Statistics 2022 - Annual Bulletin ," 2022. Accessed: Oct. 12, 2024. [Online]. Available: http://www.statistics.gov.lk/ComputerLiteracy/StaticalInformation/Bulletins/2022-Annual.

[6] G. Erdogan, R. Halvorsrud, C. Boletsis, S. Tverdal, and J. Pickering, "Cybersecurity Awareness and Capacities of SMEs," in *Proceedings of the 9th International Conference on Information Systems Security and Privacy*, 2023, pp. 296–304, doi: 10.5220/0011609600003405.

[7] M. M. Hossain, M. Fotouhi, and R. Hasan, "Towards an Analysis of Security Issues, Challenges, and Open Problems in the Internet of Things," in *2015 IEEE World Congress on Services*, Jun. 2015, pp. 21–28, doi: 10.1109/SERVICES.2015.12.

[8] A. A. Loishyn, S. Hohoniants, M. Ya.Tkach, M. H. Tyshchenko, N. M. Tarasenko, and V. S. Kyvliuk, "Development of the Concept of Cybersecurity of the Organization," *TEM J.*, pp. 1447–1453, Aug. 2021, doi: 10.18421/TEM103-57.

[9] SBA, "Stay safe from cybersecurity threats.," 2021. Accessed: Aug. 13, 2024. [Online]. Available: https://www.sba.gov/business-guide/manage-your-business/stay-safe-cybersecurity-threats.

[10] A. Chidukwani, S. Zander, and P. Koutsakis, "A Survey

on the Cyber Security of Small-to-Medium Businesses: Challenges, Research Focus and Recommendations," *IEEE Access*, vol. 10, pp. 85701–85719, 2022, doi: 10.1109/ACCESS.2022.3197899.

[11] N. Sombatruang, M. A. Sasse, and M. Baddeley, "Why do people use unsecure public wi-fi?," in *Proceedings of the 6th Workshop on Socio-Technical Aspects in Security and Trust*, Dec. 2016, pp. 61–72, doi: 10.1145/3046055.3046058.

[12] J. Firch and L. L. C. PurpleSec, ."Cybersecurity trends you can't ignore in 2021.," 2021.

[13] S. Romanosky, L. Ablon, A. Kuehn, and T. Jones, "Content analysis of cyber insurance policies: how do carriers price cyber risk?," *J. Cybersecurity*, vol. 5, no. 1, Jan. 2019, doi: 10.1093/cybsec/tyz002.

[14] A. W. Bartik, M. Bertrand, Z. Cullen, E. L. Glaeser, M. Luca, and C. Stanton, "The impact of COVID-19 on small business outcomes and expectations," *Proc. Natl. Acad. Sci.*, vol. 117, no. 30, pp. 17656–17666, Jul. 2020, doi: 10.1073/pnas.2006991117.

[15] USAID, ""ASSESSING THE DIGITAL MATURITY OF SMALL- AND MEDIUM-SIZED EXPORTERS IN SRI LANKA," Colombo, 2022.

[16] S. Wijayarathne, "Cyber Insurance in Small and Medium-Sized Enterprises in Sri Lanka," 2023.

[17] C. Cobanoglu, A. Bilgihan, K. "Khal" Nusair, and K. Berezina, "The Impact of Wi-Fi Service in Restaurants on Customers' Likelihood of Return to a Restaurant," *J. Foodserv. Bus. Res.*, vol. 15, no. 3, pp. 285–299, Jul. 2012, doi: 10.1080/15378020.2012.706194.

[18] M. A. Gregory, I. McShaneIan, and C. Wilson, "Practicing safe public wi-fi - Assessing and managing data-security risks," 2016.

[19] P. Klasnja *et al.*, "'When I am on Wi-Fi, I am fearless,'" in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, Apr. 2009, pp. 1993–2002, doi: 10.1145/1518701.1519004.

[20] T. S. Sobh, "Wi-Fi Networks Security and Accessing Control," *Int. J. Comput. Netw. Inf. Secur.*, vol. 5, no. 7, pp. 9–20, Jun. 2013, doi: 10.5815/ijcnis.2013.07.02.

[21] ENISA, "SME Cybersecurity," *European Union Agency for Cybersecurity*, 2024. https://www.enisa.europa.eu/topics/cybersecurity-education/sme_cybersecurity (accessed Oct. 12, 2024).

[22] ICTA, "Policy for Digital Transformation of Education," 2022, Accessed: Apr. 21, 2023. [Online]. Available: https://www.icta.lk/policy-for-digital-transformation-of-education/.

[23] CBSL, "Central Bank of Sri Lanka Annual Report: 2021," 2021. Accessed: Oct. 13, 2024. [Online]. Available: https://www.cbsl.gov.lk/en/publications/economic-and-financial-reports/annual-reports/annual-report-2021.

[24] ADB, "Asia SME Finance Monitor 2020: Sri Lanka," 2020. Accessed: Oct. 13, 2024. [Online]. Available: https://www.adb.org/publications/asia-sme-monitor-2020-country-regional-reviews.

Creating A Sri Lankan Meal Plan for A Diabetic Patient Using Graph Theory

R.W.K.T. Rajapaksha

NSBM Green University,
Faculty of Computing,
Sri Lanka.
kavishka.r@nsbm.ac.lk

G.H.J. Lanel
Department of Mathematics,
Faculty of Applied Sciences,
University of Sri Jayewardenepura,
Sri Lanka.
ghjlanel@sjp.ac.lk

Abstract: **Diabetes Mellitus, affecting over 420 million people globally and expected to exceed 500 million by the decade's end, is increasingly prevalent in Sri Lanka due to rising obesity and sedentary lifestyles. Approximately 11.3% of the Sri Lankan adult population suffers from diabetes, with type 2 and gestational diabetes being predominant. Effective diabetes management through diet is essential for control and prevention. This study develops culturally appropriate meal plans for type 2 and gestational diabetes using graph theory methods like vertex coloring and subgraph concepts. Data on suitable, locally available, and affordable foods were gathered from nutrition sources. Graphs representing various food items were constructed using Mathematical software, and vertex coloring identified optimal food combinations. The subgraph concept was then employed to distribute these combinations across meals. The resulting meal plans, tailored to Sri Lankan dietary preferences, offer practical and sustainable solutions for diabetes management in the local context.**

Keywords: Vertex coloring, type 2 diabetes, gestational diabetes, meal plan

I. INTRODUCTION

Diabetes Mellitus is a prevalent non-communicable disease characterized by elevated blood sugar (glucose) levels due to deficiencies in insulin production, insulin action, or both. Globally, diabetes has seen a dramatic rise, now affecting over 420 million people a figure that has quadrupled since 1980. It is projected to exceed 500 million by the end of the decade. Between 2000 and 2019, diabetes-related deaths increased by 70% worldwide. The COVID-19 pandemic further highlighted the vulnerability of diabetic patients, with a significant number experiencing severe COVID-19 complications or succumbing to the virus. In Sri Lanka, the diabetes epidemic has surged at an alarming rate, largely driven by the increasing prevalence of obesity, sedentary lifestyles, and poor dietary habits. As of 2021, 11.3% of the adult population in Sri Lanka was reported to have diabetes, underscoring the urgent need for effective management and prevention strategies.

The escalating diabetes crisis in Sri Lanka necessitates a focus on healthy eating as a key element in diabetes prevention and management. Type 2 diabetes and gestational diabetes are the most common forms seen in the country. Given the current economic challenges, it is essential to develop meal plans that incorporate affordable, locally available Sri Lankan foods that cater to the specific needs of type 2 and gestational diabetic patients. Effective meal planning plays a critical role in maintaining blood sugar levels, improving overall health, and reducing the risk of diabetes-related complications. The integration of healthy lifestyle choices, including regular exercise and a balanced diet, is fundamental to diabetes control and long-term well-being.

This research aims to address the need for culturally relevant and nutritionally appropriate meal plans for diabetic patients in Sri Lanka. By employing graph theory approaches, such as vertex coloring and subgraph concepts, this study seeks to create optimized meal plans tailored to the dietary needs of type 2 and gestational diabetic patients. This method offers a systematic approach to meal planning, ensuring that patients can access nutritious and affordable meals that align with their health requirements. Ultimately, this research contributes to the development of sustainable diabetes management practices in Sri Lanka, promoting better health outcomes for individuals affected by the disease.

II. LITERATURE REVIEW

A. The diabetes epidemic in Sri Lanka

Diabetes mellitus is on the rise over the world, posing a threat to both developed and developing nations. By 2040, the worldwide prevalence is projected to be 642 million, a 55% increase compared to 2015. Sri Lanka has not been spared from this pandemic and a similar upward trend in

prevalence has been observed in local studies [2]. Overall prevalence of diabetes for Sri Lankans aged >20 years was 10.3% according to Sri Lanka diabetes and cardiovascular study (SLDCS) conducted 10 years back in 2006 with projected prevalence of 13.9% for year 2030 [5]. Therefore, an alarming increase in complications of diabetes, both microvascular disease and macrovascular disease will be seen unless urgent measures are taken to prevent them.

B. Type 2 diabetes

Type 2 diabetes is a chronic condition in which the body does not properly use insulin, leading to insulin resistance. This type of diabetes typically affects middle-aged and older adults, although it was once primarily referred to as adult-onset diabetes. However, due to rising rates of childhood obesity, type 2 diabetes is increasingly being diagnosed in children and teenagers. Among Sri Lankans, type 2 diabetes mellitus is a prevalent condition [5]. Reducing the consumption of highly processed carbohydrates, sugary beverages, trans and saturated fats, as well as limiting red and processed meat intake, can help decrease the risk of developing type 2 diabetes [7].

C. Gestational diabetes

Gestational diabetes mellitus (GDM), which is diagnosed during pregnancy and is distinct from overt diabetes, is becoming increasingly prevalent due to the rising rates of obesity and type 2 diabetes [1]. In Sri Lanka, the incidence of GDM is also on the rise. Typically, blood sugar levels return to normal after childbirth for women who experience gestational diabetes. However, having GDM increases the likelihood of developing type 2 diabetes in the future. Although there is no definitive way to prevent gestational diabetes, adopting healthy lifestyle habits can lower the chances of recurrence in future pregnancies and reduce the risk of progressing to type 2 diabetes. Eating a diet rich in fiber, low in fat and calories, and focusing on fruits, vegetables, and whole grains can help reduce the risk of GDM [7].

D. Vertex coloring applications

A classical problem in graph theory, the graph coloring problem is to color the nodes of an undirected graph with as few colors as possible, such that no two adjacent nodes share a color [6]. Graph coloring has many applications including map coloring, task scheduling, parallel computation, network design etc. [3]. Evidence of this can be found in various papers and books, in which the coloring is studied, and the problems and conjectures associated with this field of research are being described and solved [4],[8] & [9].

Those research papers are described about the suitable meal plans for diabetic patients but not specifically mentioned the names of the foods separately for each nutrition also they are not mostly considered about the locally available healthy foods that good for diabetic. On the other hand, most of the articles are used vertex coloring approach to create scheduling for learning purposes. Therefore, this research will be carried out for creating Sri Lankan meal plans for type 2 and gestational diabetic patients using graph theory approaches including vertex coloring and subgraph concepts.

III. THEORETICAL FRAMEWORK

In graph theory, the concept of a complement graph plays a significant role in various applications. The complement of a simple graph G is defined as a graph $\acute{G}$ that contains all the vertices of G, but with edges only between the pairs of vertices that are not connected by an edge in G. Mathematically, the number of vertices in G and $\acute{G}$ are equal, and the sum of the edges in G and $\acute{G}$ equals the total number of edges in a complete graph with the same number of vertices. This relationship is represented by $|E\ (G)| + |E(\acute{G})| = |E\ (K_n)| = \frac{n\ (n-1)}{2}$, where n denotes the total number of vertices.

Vertex coloring is another important concept in graph theory, where colors are assigned to the vertices of a graph such that no two adjacent vertices share the same color. The chromatic number of a graph is the minimum number of colors required to achieve this proper coloring. Vertex coloring has practical applications in problems such as map coloring, task scheduling, and conflict resolution. Additionally, subgraphs and complete graphs are essential structures within graph theory. A subgraph is a graph formed from a subset of vertices and edges of a larger graph, and a complete graph K_n is a graph in which every vertex is connected to every other vertex. These fundamental concepts form the basis for many advanced studies and applications in graph theory.

IV. METHODOLOGY

A. Data Collection

Data on foods which are suitable for both type 2 and gestational diabetic patients are taken from the official website of the Nutrition Division of the Ministry of Health and other diabetic related nutrition websites. Foods are chosen considering the proper nutrition concepts for diabetic patients which are recommended by the doctors and other authorized parties. When select the foods have given priority for locally available Sri Lanka foods.

These are the nutrition facts that consider when select foods for type 2 diabetic patients.

- Portion control is essential
- Add more vegetables to diet
- Limit sugars and sweets (sugar, juggary, candies, jam, desserts)
- Limit fatty foods, salt
- Alcohol intake should be in moderation

These are the nutrition facts that consider when select foods for gestational diabetic patients.

- Plenty of whole fruits and vegetables
- Moderate amounts of lean proteins and healthy fats
- Moderate amounts of whole grains, such as bread, cereal, pasta, and rice, plus starchy vegetables, such as peas
- Fewer foods that have a lot of sugar, such as soft drinks, fruit juices, and pastries
- Regular intake of products containing folic acid and vitamin A
- Sweet fruits should be limited to one meal per week. Should also avoid frequent use of avocados

B. Creating Graphs

Considering the selected suitable foods as vertices and the edges as connections between different nutritional foods, two graphs are created using Maple software. To apply the vertex coloring approach, it is necessary to obtain the complement of both graphs. (Figure 1 and Figure 2)

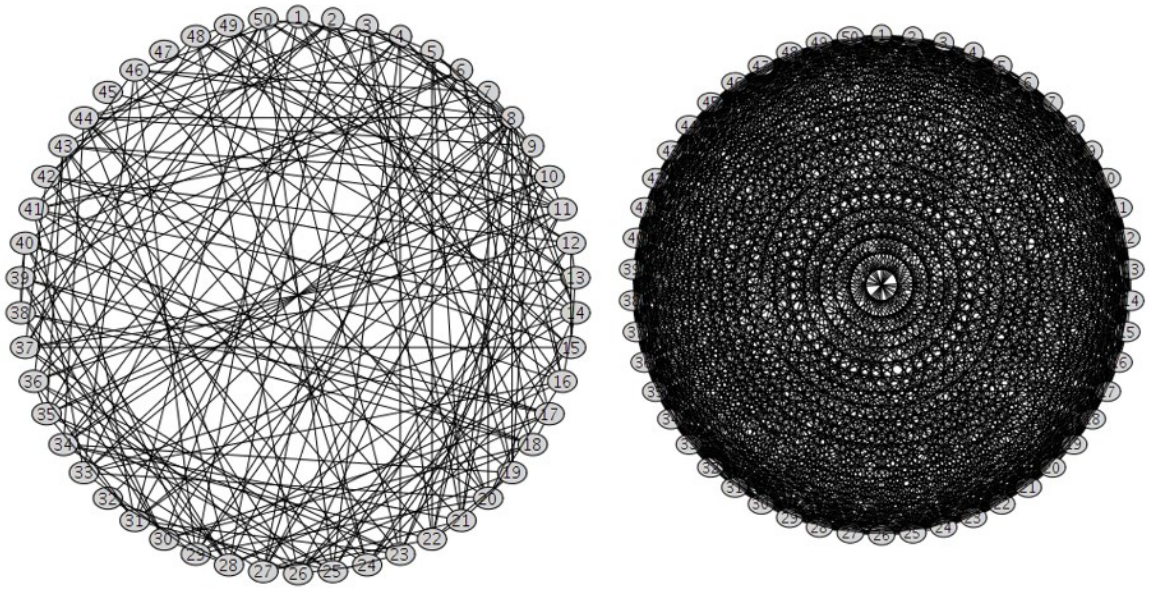


Figure 1: Graph for type 2 diabetic and its complement graph

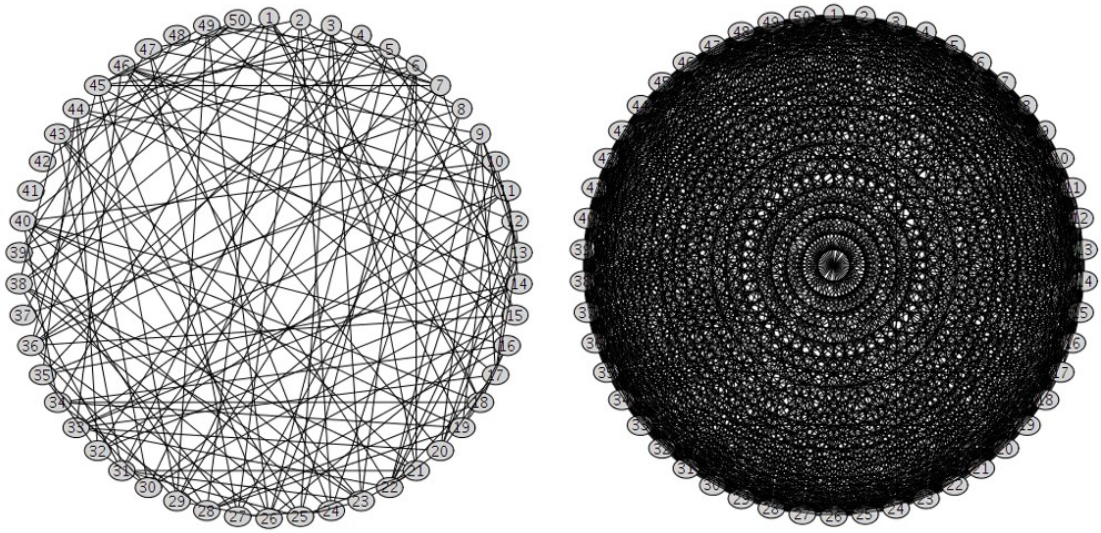


Figure 2: Graph for gestational diabetic and its complement graph

C. *Applying vertex coloring and sub graph concepts*

The vertex coloring approach is then applied to these complement graphs. The chromatic number, which indicates the minimum number of colors needed to properly color the graph, is determined for each complement graph. This chromatic number helps in grouping the foods into distinct sets where no two adjacent foods have the same color. (Figure 3and Figure 4)

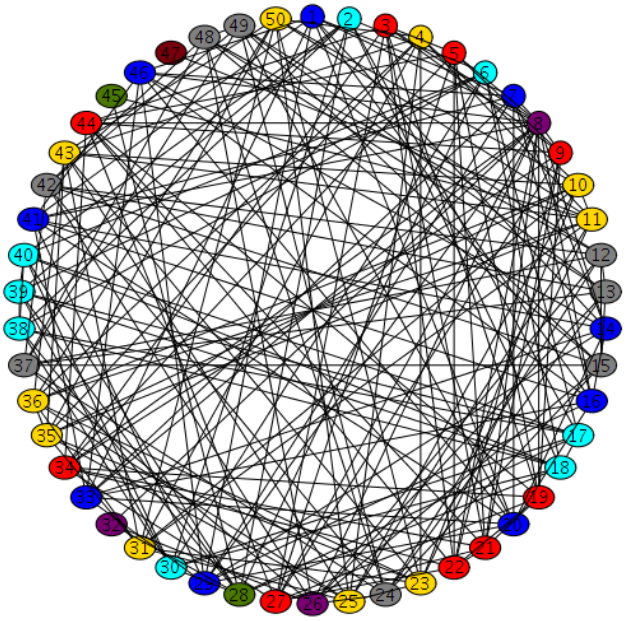


Figure 3: Vertex coloring for graph of type 2 diabetes of gestational diabetes

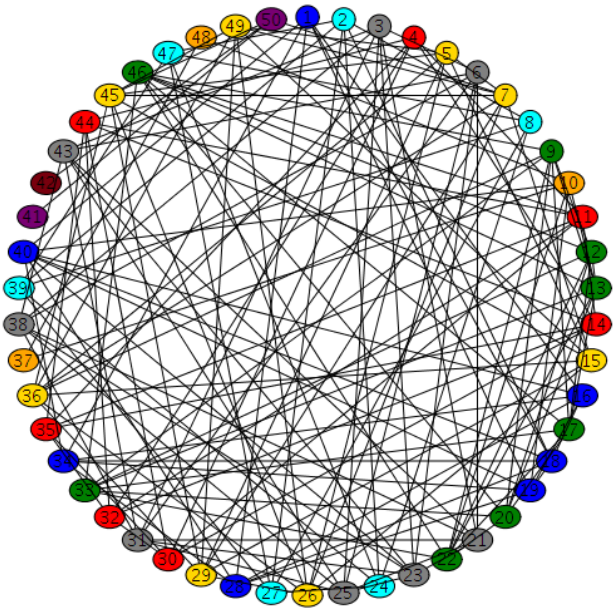


Figure 4: Vertex coloring for graph

D. Creating Meal Plans

Based on the chromatic number, the corresponding vertex sets are identified. These vertex sets are then used to color the vertices in the main graphs. Among the vertex sets, the most appropriate sets are selected to create meal plans. Finally, the subgraph concept is applied to allocate each vertex set to specific meals; breakfast, lunch, and dinner. The nutritional guidelines for diabetic patients are considered while forming these subgraphs.

For Type 2 Diabetic Patients:

- Breakfast: Foods that are high in fat and protein, low in carbohydrates, and rich in fiber.
- Lunch: Low fiber carbohydrates (e.g., rice, sweet potatoes), non-starchy vegetables (e.g., green leaves, broccoli), high fiber vegetables (e.g., beans, whole grains), lean protein (e.g., tuna, eggs, or chicken once a week), and fruits (e.g., apples).
- Dinner: Non-starchy vegetables (e.g., green beans, broccoli, cauliflower, cabbage, and carrots), lean protein, carbohydrate-rich foods (e.g., grains, starchy vegetables like potatoes), rotti, string hoppers, and low-calorie drinks.

For Gestational Diabetic Patients:

- Breakfast: Low fat, lean protein, carbohydrates, high fiber foods, and starchy vegetables such as peas.
- Lunch: Foods rich in vitamin A, folic acid, lean protein, less sweet fruits, fiber, and minerals.
- Dinner: Vegetables, low fat, mineral-rich foods (e.g., grains and wheats), and lower carbohydrate options.

V. RESULTS AND DISCUSSION

For type 2 diabetes, the graph's chromatic number was determined to be 8, which facilitated the creation of 5 distinct vertex sets. Each vertex set represents a unique combination of nutritional components, enabling the allocation of diverse foods across breakfast, lunch, and dinner for each day. Red rice (with outer hull) was not included as a vertex due to its necessity in daily meals; thus, it was treated as a constant component (Table 1). This approach allowed for the development of well-balanced meal plans that patients can repeat cyclically, ensuring consistency in nutrition while adhering to dietary restrictions.

Similarly, for gestational diabetes, the chromatic number of the graph was found to be 9, which permitted the formation of 6 distinct vertex sets. These sets were used to systematically allocate foods for breakfast, lunch, and dinner, ensuring a varied and nutritionally adequate meal plan (Table 2). By incorporating the subgraph concept, the methodology provided a structured framework for meal planning that accommodates the unique dietary needs of gestational diabetes patients. As with type 2 diabetes, red rice was considered a fixed component in the meal plans, reinforcing its role as a staple food.

Table 1: Five days meal plans for type 2 diabetic patients

	Breakfast	**Lunch**	**Dinner**
Day 1	Peas, Herbel Porridge	Red rice with outer hull, Ambarella, Sweet potato, Mackerel tuna, Tylophora pauciflora, Yogurt	Red rice flour roti, Mackerel tuna, orange juice
Day 2	Red rice with outer hull, Sardine, Spinach, Pumpkin, Coffee without sugar	Heenati rice, Banana blossom, Spinach, Sardine, Watermelon	Red rice String hoppers, Potato curry
Day 3	Red rice with outer hull, Bitter guard, Joyweed, Salmon fish	Suwandel rice, Bitter guard, Lady's fingers, Joyweed, Salmon fish, Apple	Whole meal bread, Mushroom soup
Day 4	Green gram, Oatmeal	Red rice with outer hull, Banlkarela, Marrow, Mushroom, Passionfruit leaves, Avocado	Carrot, Tuna
Day 5	Manioc	Red rice with outer hull, Beans, Cauli flower, Thembu leaves, Egg, Star fruit	Vegetable soup, Whole grain crackers

Table 2: Six days meal plans for gestational diabetic patients

	Breakfast	**Lunch**	**Dinner**
Day 1	Green gram	Red rice with outer hull, Banana, Pumpkin, Beans, Chicken	Low fat cheese, Vegetable salad, Parsley
Day 2	Peas	Red rice with outer hull, Tuna, Spinach, Ambarella, Green apple	Whole grain bread, Peanut butter
Day 3	Sweet potato, orange juice	Red rice with outer hull, Egg, Cauliflower, Tylophora pauciflora, Yogurt	Whole wheat pasta
Day 4	Herbal porridge, red rice flour rotti	Red rice with outer hull, Salmon, Grapes, Eggplant, Cucumber	Apple juice without sugar, Whole grain crackers
Day 5	Red rice flour pittu, Beef liver curry	Red rice with outer hull, White cabbage, Mango, Carrot, Beef liver	Chicken-vegetable soup
Day 6	Oats, Low fat milk	Red rice with outer hull, Tilapia, Mushroom, Pears, Broccoli	Red rice flour string hoppers, Dal curry

CONCLUSION

The prevalence of diabetes, particularly type 2 and gestational diabetes, has risen significantly in Sri Lanka due to unhealthy lifestyle choices, such as poor diet and lack of physical activity. The need for effective meal planning tailored to the nutritional needs of diabetic patients is therefore critical. This research aimed to create meal plans specifically for type 2 and gestational diabetic patients by utilizing graph theory techniques, such as vertex coloring and subgraph concepts.

By focusing on locally available foods that are affordable and accessible in Sri Lanka, this study provides practical and sustainable meal plans that incorporate appropriate nutrition guidelines for diabetic patients. Through the application of graph theory, the research successfully identified food combinations that meet the nutritional requirements for breakfast, lunch, and dinner, ensuring variety and balance in the diet while adhering to diabetic restrictions.

The methodology employed allowed for the systematic allocation of foods into meal plans, considering the chromatic number to ensure a proper distribution of nutrients across the meals. This approach offers a novel contribution to diabetic meal planning by integrating mathematical concepts with nutritional science, leading to the development of customized and efficient meal plans for type 2 and gestational diabetes patients in Sri Lanka.

Overall, this research not only provides a framework for creating diabetic-friendly meal plans but also emphasizes the importance of incorporating local, healthy foods into the daily diet of diabetic patients. The findings can serve as a basis for further research and practical application in diabetes management, contributing to better health outcomes for those affected by the disease.

REFERENCES

[1] Coustan, DR 2013, Gestational diabetes mellitus, *Clinical chemistry*, *59*(9), pp.1310-1321.

[2] *Diabetes mellitus management guidelines*, Sri Lanka College of Endocrinologists, 2018.

[3] Dandashi, A and Al-Mouhamed, M 2010, May, 'Graph coloring for class scheduling', In *ACS/IEEE International Conference on Computer Systems and Applications-AICCSA,* (pp. 1-4). IEEE.

[4] Formanowicz, P and Tanaś, K 2012, 'A survey of graph coloring-its types, methods and applications', *Foundations of Computing and Decision Sciences*, *37*(3), pp.223-238.

[5] Katulanda, P, Constantine, GR, Mahesh, JG, Sheriff, R, Seneviratne, RDA and Wijeratne, S 2008 'Prevalence and projections of diabetes and pre-diabetes in adults in Sri Lanka—Sri Lanka Diabetes', Cardiovascular Study (SLDCS) Diabet Med, 25, pp. 1062– 1069.

[6] Klotz, W 2002, 'Graph Coloring Algorithms', Mathematic- Bericht 5, TU Clausthal.

[7] Shanmuganathan, Y and Hettiarchchi, J 2016, '*Food Exchange List for Sri Lankans with Type 2 Diabetes Mellitus*', Faculty of Livestock, Fisheries and Nutrition, Wayamba University of Sri Lanka. Makandura.

[8] Toft, B, Jensen, TR 1995,'Graph Coloring Problems', John Wiley and Sons, New York, 1995.

[9] Toft, B, Jensen, TR 2001, '25 pretty graph colouring problems', Discrete Mathematics, 229:167169.

Improving Efficiency in CI/CD Processes through Cloud-Native Automation and Container Deployment

SNG Fernando
Fortude (Pvt) Ltd
nikilaf@fortude.co

Chamara Disanayake
Faculty of Computing
NSBM Green University
Pitipana, Homagama, Sri Lanka
chamara.d@nsbm.ac.lk

Chamindra Attanayake
Faculty of Computing
NSBM Green University
Pitipana, Homagama, Sri Lanka
chamindra.a@nsbm.ac.lk

***Abstract*— The rapid evolution of software development underscores the critical role of Continuous Integration and Continuous Deployment (CI/CD) processes. This research delves into the intricate nuances of current CI/CD practices, identifying prevalent inefficiencies primarily resulting from the fragmented use of various tools coupled with substantial manual human intervention. Such practices often lead to significant delays, posing challenges like protracted pipeline constructions and time-intensive tasks such as Docker file authoring. To address these inefficiencies, the study introduces an innovative solution, a Command-Line Interface (CLI) application designed using the Go programming language. The proposed system seeks to optimize CI/CD operations by embracing cutting-edge cloud-native architectures by seamlessly automating crucial stages. From code building and scanning to Docker image generation and deployment, the application aims to reduce manual touchpoints, thus streamlining operations and markedly enhancing efficiency. Grounded in a combination of comprehensive literature reviews, organizational data collation, and empirical testing, this research offers a deeper understanding of the current challenges in CI/CD and presents a forward-looking perspective, charting a strategic course for further academic inquiries and industry-focused advancements.**

Keywords—DevOps, CICD, Golang, Docker

I. INTRODUCTION

Software development methodologies have undergone significant transformations, with Continuous Integration/Continuous Deployment (CI/CD) processes emerging as pivotal elements in the modern DevOps landscape. CI/CD practices facilitate the frequent integration of code changes and enable rapid delivery of software to production, ensuring a streamlined development lifecycle and enhanced product quality[1], [2]. Despite the widespread adoption of these methodologies, current CI/CD practices encounter substantial challenges. Manual interventions, difficulties in integrating disparate tools, and the complexities of managing intricate CI/CD pipelines hinder the achievement of full automation and efficiency, particularly in cloud-native environments and container deployments [3]

This research aims to address the inefficiencies inherent in CI/CD processes by exploring the potential of cloud-native automation and container deployment to streamline these practices. Specifically, the study evaluates and validates approaches, tools, and practices in continuous integration and deployment, leveraging insights from a real-world software development organization. By adopting a cloud-native approach and employing a Go language-written script for automating the CI/CD process—including code build, Docker build, code scan, and artifact publishing—this research seeks to mitigate the identified challenges. The goal is to enhance CI/CD processes, particularly in terms of deployment time and automation in cloud-native settings, thus contributing significantly to the field of software development [4].

Through this investigation, the thesis titled "Improving Efficiency in CI/CD Processes through Cloud-Native Automation and Container Deployment" aspires to illuminate the critical role of CI/CD in the evolution of DevOps practices. It also aims to address the pressing issues of manual interventions, tool integration difficulties, and the complexity of managing CI/CD pipelines, thereby paving the way for more efficient, automated software development cycles.

II. PROBLEM STATEMENT

CI/CD processes have revolutionized software development by enabling frequent code integration and rapid software delivery. However, current CI/CD practices face challenges due to manual intervention, tool integration issues, and managing complex CI/CD pipelines, leading to inefficiencies and reduced software quality. This research aims to address these challenges by developing a solution that automates CI/CD processes and effectively integrates cloud-native automation and container deployment. The goal is to help software development teams deliver high-quality software more quickly and reliably, improving productivity and customer satisfaction.

Key Points:

- Current CI/CD processes are inefficient due to manual intervention, tool integration issues, and complex pipelines.
- Manual CI/CD processes lead to delays, increased costs, and reduced product quality.
- Cloud-native environments and container deployment add further complexities to CI/CD processes.
- This research seeks to automate CI/CD processes and effectively integrate cloud-native automation and container deployment.

Ultimate Goal:

- Develop a solution that enhances the effectiveness and efficiency of CI/CD processes for cloud-native applications.
- Enable software development teams to deliver high-quality software more quickly and reliably.
- Improve productivity and customer satisfaction in the software development lifecycle.

Table 1 shows dataset summarized average durations for different stages of manual code deployment in a software development organization. This analysis, based on 50 code builds, excludes test case times and focuses on the manual aspects of deployment. It includes time spent on writing code, conducting scan processes, building Docker images, and the overall time from code build to Docker image completion. Emphasis is on the manual effort involved in writing pipeline scripts, managing pipelines, and creating Docker files for various programming languages, highlighting the time investment in these tasks.

Table 3

Programming Language	**Average Build Time (Manual)**	**Average Scan Time (Manual)**	**Average Docker Image Build**	**Total Time from Code**
			Time	**Build to Docker Image Build**
Java (Spring Boot)	30 minutes	20 minutes	10 minutes	70 minutes
Python	25 minutes	15 minutes	8 minutes	60 minutes
JavaScript	28 minutes	18 minutes	9 minutes	65 minutes

a. Reseach Question and Motivation

Research Question: This study investigates the feasibility of leveraging cloud-native automation and container deployment to overcome challenges in Continuous Integration and Continuous Deployment (CI/CD) processes. The research aims to design and implement a comprehensive solution addressing obstacles such as manual intervention, tool integration issues, and pipeline complexity. The central question guiding the investigation is: "How can we design and implement a comprehensive solution that leverages the power of cloud-native automation and container deployment to significantly enhance the efficiency of CI/CD processes, thereby reducing the time taken for code build, scan, and Docker image build processes, and ultimately improving the productivity and satisfaction of software development teams?"

Research Motivation: The motivation for this research arises from the critical role CI/CD processes play in modern software development. While these processes offer substantial benefits, practical challenges persist, particularly in cloud-native environments and container deployment. Manual intervention, tool integration issues, and pipeline complexity can hinder efficiency, leading to delays, increased costs, and reduced product quality. The study is driven by the belief that leveraging cloud-native automation and containers can substantially improve CI/CD efficiency, ultimately enabling software development teams to deliver high-quality software more quickly and reliably. The broader motivation lies in the potential impact on the software industry, where enhanced CI/CD efficiency can lead to reduced time to market, improved product quality, increased competitiveness, and overall success in the marketplace.

Key Points

1. Current Challenges in CI/CD Implementation:
    - Manual intervention issues
    - Lack of seamless tool integration
    - Complexity in managing and scaling CI/CD pipelines

2. Significance of CI/CD in Modern Software Development:
    - Role of CI/CD in enabling frequent integration of changes
    - Swift and reliable software delivery for maintaining competitiveness
3. Motivation for Leveraging Cloud-Native Automation and Containers:
    - Potential benefits of cloud-native automation
    - Impact of container deployment on CI/CD efficiency
    - Time-to-market reduction
    - Improved product quality and customer satisfaction
    - Increased competitiveness in the software industry

III. LITERATURE REVIEW

b. Conceptual Taxonomy of the Literature Organization

This literature review adopts a conceptual taxonomy that systematically organizes the themes and topics related to Continuous Integration and Continuous Delivery (CI/CD) processes, cloud-native automation, and container deployment. This framework ensures a comprehensive and coherent exploration of the existing body of knowledge, guiding the analysis and synthesis of relevant research [1].

B. Domain Overview

This research delves into the intersection of software development, cloud-native automation, and container deployment, specifically within the context of CI/CD processes. It places particular emphasis on the stages of code building, Docker image creation, and code scanning within the CI/CD pipeline [2].

C. Software Engineering

Modern software engineering practices have witnessed a paradigm shift, embracing methodologies such as Agile and DevOps, which prioritize frequent software integration and delivery [3]. This has led to the widespread adoption of CI/CD processes as cornerstones of contemporary software development. CI/CD processes enable developers to integrate their changes into a shared repository frequently, fostering early problem detection and ensuring swift, reliable software delivery [4]. However, practical implementation of CI/CD processes poses challenges, including:

- Complexity in managing and coordinating multiple stages of the pipeline [5]
- Need for robust testing and quality assurance mechanisms [3]
- Seamless integration of various tools and platforms [4]

D. DevOps Practices and Modern Software Engineering

DevOps practices have gained significant momentum in modern software development, aiming to bridge the gap between development and operations teams [3]. This collaborative approach fosters a culture of shared responsibility, leading to faster development cycles, improved deployment quality, and more reliable software [5]. Key principles of DevOps include:

- Automation of processes, including integration, testing, deployment, and monitoring [3]
- Infrastructure as Code (IaC), enabling automated provisioning and management of IT infrastructure [6]
- Continuous learning and improvement, emphasizing experimentation and adaptation [3]

E. CI/CD Processes

CI/CD processes have emerged as pivotal mechanisms for efficient and reliable software delivery. By facilitating frequent integration of changes into a shared repository, they enable early problem detection and swift resolution, ensuring rapid and dependable software updates [4]. However, practical implementation of CI/CD processes encounters challenges, such as:

- Lack of comprehensive automation, often requiring significant manual intervention [5]
- Complexities in managing and scaling CI/CD pipelines, particularly in cloud-

native environments and container deployment [7]

F. Cloud-Native Automation

Cloud-native automation has revolutionized software development by automating application lifecycle processes from development to deployment and management, leveraging the inherent scalability, flexibility, and resilience of cloud environments [6]. It underpins contemporary software development practices, including DevOps and CI/CD, enabling seamless integration, continuous delivery, and rapid deployment of applications [7].

G. Challenges in Implementing Cloud-Native Automation

Despite its transformative potential, cloud-native automation presents challenges, including:

- Managing and orchestrating services across diverse cloud environments [8]
- Implementing robust security measures to protect sensitive data and applications [9]
- Addressing the skills gap in cloud-native technologies [10]

H. Container Deployment

Container deployment is another critical aspect of contemporary software development, offering a lightweight and efficient approach to packaging and deploying applications [2]. Essentially, a container is a self-contained unit of software, complete with its code, runtime, system tools, libraries, and settings. This encapsulation ensures that the software runs reliably and consistently regardless of the environment it's deployed in.

IV. RESEARCH METHODOLOGY

This chapter delves into the research methodology employed to study the complexities and challenges of Continuous Integration and Continuous Deployment (CI/CD) processes in the evolving landscape of software development and IT operations. With cloud-native solutions and DevOps principles gaining traction, optimizing CI/CD processes becomes crucial. This chapter serves as a roadmap, outlining the research methods for exploring these processes and proposing practical solutions.

This section describes the research design for developing and evaluating a CI/CD automation tool to address the limitations of traditional CI/CD processes. These processes, while facilitating early issue detection and rapid software delivery, often face practical challenges hindering efficiency.

A . Nature of the Research

Exploratory: We delve into the challenges hindering CI/CD efficiency, including manual intervention, lack of seamless tool integration, and complex CI/CD pipeline management, particularly in cloud-native environments and container deployments. These issues can lead to delays, increased costs, and diminished product quality.

Applied: We design and implement a practical CLI tool to address these challenges. This tool aims to streamline containerization, improve developer-DevOps collaboration, and enhance overall CI/CD pipeline efficiency.

B Methodological Framework.

- Tool Development using Go (Golang): The tool was developed using Go due to its robustness and efficiency for scalable and high-performance *systems.*
- Automatic Code Detection: The tool is to automatically detects the programming language used in a code repository, aiding in appropriate environment and dependency encapsulation during containerization.
- Builder Concept: We introduce the concept of pre-defined builders – templates or configurations for various programming languages and frameworks. The corresponding builder will be invoked for containerization upon code detection, eliminating manual Docker-file creation.
- Automated Image Tagging and Repository Push: The tool will automatically tag the generated image using a predefined naming convention and push it to a designated container registry.
- Static Code Vulnerability Scanning: The tool will conduct static code scans on the container image, identifying potential vulnerabilities and areas for code improvement. This emphasizes the importance of code scanning within the CI/CD pipeline for code quality and security.
- Comprehensive Report Generation: Following the scan, a

detailed report will be generated, encompassing code quality metrics and identified vulnerabilities. This report will serve as a valuable resource for developers to refine their codebase.

Expected Outcomes

This research design aims to develop a comprehensive CI/CD automation tool that addresses the challenges of traditional CI/CD processes. By integrating features like automatic code detection, builders, and vulnerability scanning, the tool aspires to streamline the software development lifecycle.

V. FUNCTIONAL REQUIREMENTS:

The system's functional requirements entail facilitating repository interaction, allowing users to either clone or select pre-cloned repositories, and incorporating the capability to identify and prompt for the programming language. The application is mandated to construct container images, intelligently choosing an appropriate builder based on the detected language, and allowing users to assign tags to these images. Anticipated future functionalities encompass vulnerability scanning, issuance of alerts for identified issues, pushing images to a registry, and providing an option for workspace cleaning. The non-functional requirements underscore the significance of usability by providing a user-friendly interface, lucid messaging, and succinct prompts. Performance criteria include expeditious language detection and image construction, streamlined future vulnerability scans, and dependable repository cloning and image-building processes. The system is expected to handle errors gracefully, demonstrate maintainability through well-organized code and comprehensive documentation, and ensure portability across diverse platforms. Scalability is imperative, enabling the system to accommodate larger repositories and simultaneous builds as it evolves. Security considerations encompass the utilization of up-to-date vulnerability scanning definitions, the preservation of confidentiality for processed data, and prohibiting external storage or transmitting repository data without explicit user consent, aligning with IEEE standards and scholarly conventions.

Proposed workflow of the system.

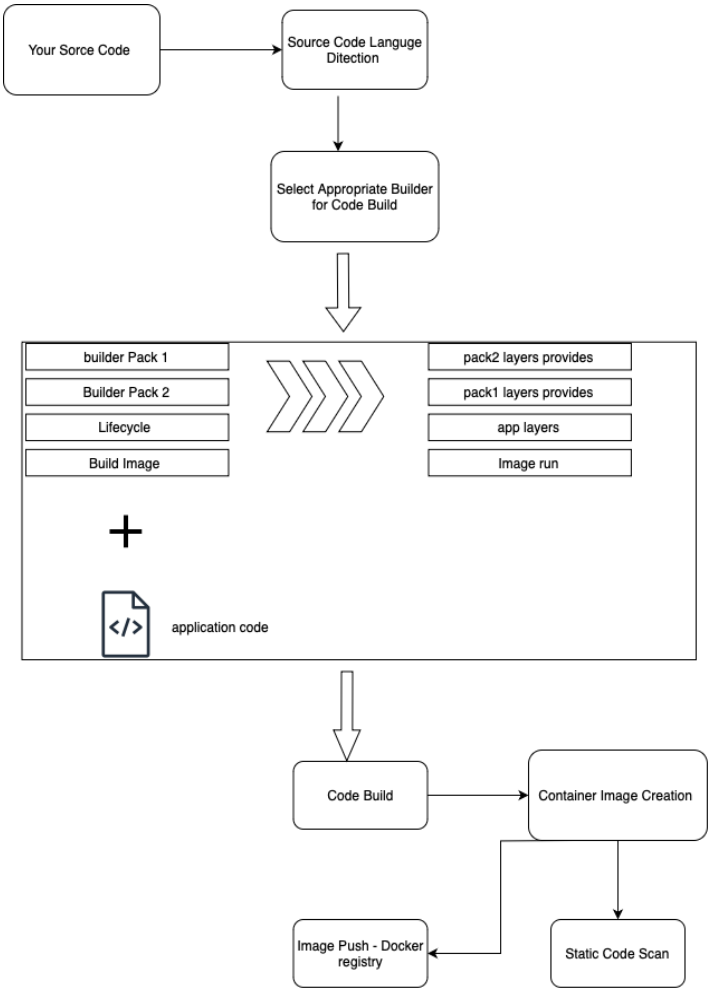


Fig. 1. System Workflow

VI. TESTING AND EVALUATION

c. Ensuring Reliability and Efficiency in a Cross-Platform CLI Application:

Section VI meticulously details the testing and evaluation processes employed to assess the performance and functionality of the developed CLI application. The testing approach encompasses module-wise scrutiny, isolating Language Detection, Image Builder, Image Scanner, Image Pusher, and the CLI Application module for individual assessment. Following module-specific tests, the application undergoes comprehensive evaluation as a unified CLI tool, emphasizing memory usage, execution speed, and adaptability across diverse environments. The testing spans Linux and Windows environments, leveraging the platform independence of Go packages and aligning with IEEE standards.

The Language Detection Module undergoes thorough testing to validate its accuracy and reliability in identifying the programming language of repositories.

The module demonstrates commendable precision in single-language scenarios but faces challenges in mixed-language environments. A manual language specification feature is introduced to address this, ensuring user control and mitigating the impact of automated detection limitations. The testing methodology involves multiple repositories representing five primary languages, revealing the module's strengths and limitations. The Go language and Docker integration efficiency is highlighted for cross-platform testing.

d. Main Module Testing – Image Builder

The Image Builder Module is then scrutinized for its pivotal role in constructing Docker images from source code. Challenges include Maven dependencies in Java, leading to frequent anomalies, and redundant storage consumption. Solutions involve introducing self-updating mechanisms for dynamic library ecosystems and adopting image pull policies to minimize storage wastage. The module, despite challenges, is optimized to handle diverse languages and dependencies, showcasing iterative improvements through rigorous testing. Full application testing encompasses various operating environments, revealing challenges like version mismatches in Go, resolved through proactive mechanisms. The chapter concludes with a summary highlighting the commitment to delivering a reliable and efficient CLI application, with insights from testing informing future iterations.

VII. CONCLUTION AND REMARKS

The research successfully achieved its objectives by identifying and analyzing challenges in CI/CD processes through a meticulous exploration rooted in comprehensive data collection methods and qualitative fact-gathering. This involved direct engagement with organizational practices, structured interviews with key stakeholders, and an immersive approach to understand the real-world application and challenges of CI/CD operations. The dual-pronged strategy provided a holistic understanding, forming the foundation for subsequent design and implementation phases. The research connected academic study with practical challenges, offering a well-rounded perspective.

A comprehensive solution was designed to address challenges, resulting in the creation of a Command Line Interface (CLI) application in the Go programming language. The CLI application streamlined and automated key aspects of the CI/CD workflow, showcasing notable speed enhancements and process optimization. The utilization of Go demonstrated its versatility, particularly in synergy with cloud-native architectures, fostering continuous learning and operational insights.

e. Challenges Encountered

The research journey encountered several challenges, contributing to a rich academic experience. Adapting to Go, a programming language with concurrent execution models and unique syntax, required a shift in programming mindset. Understanding cloud-native architectures, specifically the 'Modular Monolithic' paradigm, demanded a delicate balance between simplicity and modularity. Delving into container image architecture and builder concepts involved navigating intricate layers, dependencies, and runtime environments. The language detection module faced technical intricacies, especially in the context of multi-language repositories, requiring a balance between speed and accuracy. Data collection limitations within a specific organizational context constrained the scope of insights.

VIII. FUTURE RECOMMENDATIONS

The culmination of this research journey reveals the expansive and continually evolving nature of CI/CD, cloud-native automation, and container deployment. While the developed CLI application addresses current challenges, it also points to areas for future refinement and exploration. Recommendations include:

- Modularity and Extensibility: Envision a more modular CLI architecture for easy integration of additional functionalities, allowing developers to adapt to evolving CI/CD paradigms.
- Integration with Contemporary Tools: Ensure the CLI application's adaptability to emerging tools and methodologies, fortifying its relevance in dynamic tech ecosystems.
- Enhanced Monitoring and Logging: Integrate sophisticated monitoring and logging mechanisms to gain granular insights into the CI/CD process, enhancing anomaly detection and streamlining troubleshooting.

- Scalability Solutions: Address the imperative need for graceful scalability by considering distributed computing paradigms or integration with tools prioritizing scalability.
- Security Augmentation: Prioritize fortifying the CLI application against cyber threats by securing communications, conducting rigorous code scans, and safeguarding Docker image deployments.
- User Experience Refinements: Focus on refining the tool's intuitiveness and explore the potential for a complementary Graphical User Interface (GUI) to broaden accessibility.

REFERENCES

[1] A. Kumbhar, M. Shailaja, and R. Shankar Anupindi, "GETTING STARTED WITH CONTINUOUS INTEGRATION IN SOFTWARE DEVELOPMENT," 2018.

[2] I. Lokaadinugroho, "DevOps CICD in Higher Education," 2022, doi: 10.21203/rs.3.rs-1500352/v2.

[3] M. Senapathi, J. Buchan, and H. Osman, "DevOps capabilities, practices, and challenges: Insights from a case study," in ACM International Conference Proceeding Series, Association for Computing Machinery, Jun. 2018. doi: 10.1145/3210459.3210465.

[4] B. Laster, Jenkins 2: Up and Running, United States of America: O'Reilly Media, Inc., 2018.

[5] F. Zampetti, S. Geremia, G. Bavota, and M. Di Penta, "CI/CD Pipelines Evolution and Restructuring: A Qualitative and Quantitative Study." [Online]. Available: https://madnight.github.io/githut/

[6] N. Pathania, Beginning Jenkins Blue Ocean. Apress, 2019. doi: 10.1007/978-1-4842-4158-5.

[7] S. F. Rayanagoudar, P. S. Hampannavar, J. D. Pujari, and V. K. Parvati, "Enhancement of CICD Pipelines with Jenkins BlueOcean," International Journal of Computer Sciences and Engineering, vol. 6, no. 6, pp. 1048–1053, Jun. 2018, doi: 10.26438/ijcse/v6i6.10481053.

[8] M. K. Abhishek, D. R. Rao, and K. Subrahmanyam, "Framework to Deploy Containers using Kubernetes and CI/CD Pipeline," International Journal of Advanced Computer Science and Applications, vol. 13, no. 4, pp. 522–526, 2022, doi: 10.14569/IJACSA.2022.0130460.

[9] V. Armenise, "Continuous delivery with Jenkins: Jenkins solutions to implement continuous delivery," in Proceedings - 3rd International Workshop on Release Engineering, RELENG 2015, Institute of Electrical and Electronics Engineers Inc., 2015, pp. 24–27. doi: 10.1109/RELENG.2015.19.

[10] A. M. Buttar et al., "Optimization of DevOps Transformation for Cloud-Based Applications," Electronics (Switzerland), vol. 12, no. 2, Jan. 2023, doi: 10.3390/electronics12020357.

On-site Web Vulnerability Scanner for Agile Development Environment

Sinali Hengedata
Dept. of Computer Security and Network Systems
NSBM Green University
Homagama, Sri Lanka
sinalihengedara2020@gmail.com

Chamara Disanayake
Dept. of Computer Security and Network Systems
NSBM Green University
Homagama, Sri Lanka
chamara.d@nsbm.ac.lk

Madusanka Mithrananda
Dept. of Computer Security and Network Systems
NSBM Green University
Homagama, Sri Lanka
madusanka.m@nsbm.ac.lk

***Abstract*— In the contemporary digital landscape, websites are essential to online activities, utilizing technologies such as HTML, CSS, and JavaScript. However, the increasing complexity of web applications elevates their susceptibility to security vulnerabilities like SQL injection and Cross-Site Scripting (XSS), which can result in significant data breaches, service disruptions, and exploitation of sensitive information. This paper presents an onsite vulnerability scanner specifically designed for web developers to detect these vulnerabilities before deployment. Developed using Agile methodology, the scanner utilizes a robust technology stack, including Python for backend operations and the MEAN stack (MongoDB, Express.js, Angular, Node.js) for an intuitive user interface. The scanner's value lies in its ability to proactively identify and report common yet critical vulnerabilities, thereby enabling developers to address potential security issues early in the development process. Additionally, the tool incorporates features for user feedback and inquiries, supporting continuous improvement and adaptability. By integrating this scanner into their development workflow, developers can significantly enhance the security of their web applications, protecting against prevalent and potentially devastating security threats.**

Keywords*— *SQL injections, Cross Site Scripting, Vulnerability scanner, Web security, Websites

I. INTRODUCTION

a. Introduction

Websites and web applications are an integral part of modern life, but their security against web-based attacks is a significant concern. The difference between websites and web applications depends on their functionality and their design complexity. The purpose of a website is to publish information for users via a platform and web applications are developed to perform tasks by engaging with the users. In this case, web applications are more in danger due to their engagement with the users [1]. In today's digital age, almost every aspect of daily life has transitioned to virtual platforms. Activities such as shopping, banking, and entertainment are now commonly conducted online, offering convenience over traditional manual methods. However, this digital shift comes with its own set of challenges, notably the vulnerability of websites to web-based attacks.

Web applications are crucial for individuals as they provide seamless access to services and information, enhancing everyday activities. For companies and organizations, web applications represent their digital fingerprint, serving as the primary interface for customer interaction, service delivery, and brand presence. As such, the security of these applications is paramount, not only to protect sensitive data but also to maintain trust and credibility.

Companies invest significant resources in securing their web applications to prevent breaches that could lead to financial losses, reputational damage, and legal repercussions. Despite these efforts, ensuring robust security remains a formidable challenge. Web-based attacks can be defined as malicious activities aimed at compromising the integrity or security of web applications, websites, or web servers. According to a blog on Bright Security website [2], Brute Force Attack, Injection Attacks, Fuzz Testing (Fuzzing), Cross-Site Scripting (XSS), DDoS (Distributed Denial-of-Service), Cross-Site Request Forgery (CSRF) are some of the common web-based attacks.

For developers, building secure web applications is a complex task. They must constantly stay updated with evolving threats and implement rigorous security measures throughout the development process. This includes secure coding practices, regular security testing, adherence to industry standards and security frameworks that categorize the security requirements for web applications based on the criticality of operations and data [3]. However, the dynamic nature of web development, coupled with the pressure to deliver applications quickly, often makes it difficult to ensure comprehensive security.

After conducting thorough research, it became evident that creating a simple on-site vulnerability scanner could significantly mitigate the risk of web-based attacks. Such a scanner would enable web development teams to test their applications and websites for these vulnerabilities prior to deployment, ultimately providing the public with more secure online experiences.

b. Problem Identification

Developers face significant challenges in ensuring the security of their web applications. Firstly, it is hard to manually check each component for vulnerabilities due to the sheer volume of code and the intricate nature of modern web applications. The process requires a deep understanding of various security threats, which is not always feasible for every developer.

Moreover, the complexity of web applications adds another layer of difficulty. Modern applications often involve multiple frameworks, libraries, and third-party services, each of which can introduce potential security flaws. Identifying and mitigating these vulnerabilities requires specialized knowledge and tools, making the task even more daunting. Also using more sophisticated web vulnerability assessment tools require some knowledge and expertise that is lacking in most of the web developers.

The cost of thoroughly checking every piece of the application can be prohibitive. Investing in dedicated security experts, sophisticated scanning tools, and comprehensive testing processes demands considerable financial resources. For small-scale software companies, this expense can be particularly burdensome, potentially diverting funds from other crucial areas of development.

According to the OWASP Top 10 [4], injection attacks are considered highly common and critical, ranking third on the list. SQL injections involve injecting malicious SQL code into a backend database to gain unauthorized access to data. Cross-Site Scripting (XSS) attacks occur when malicious scripts are injected from the client's side into web pages viewable by others.

SQL injections target web applications, leading to issues such as identity theft, violations of data integrity, confidentiality, availability, and unauthorized administrative control of the server. Conversely, XSS attacks target users rather than the application itself. These attacks can result in client machines or devices becoming infected with malware, exacerbating the situation.

Attackers may have various motivations for conducting these attacks, including financial gain, espionage, political reasons, or the intent to damage the reputation of individuals or organizations. Regardless of the attack type, the societal impact is significant. Therefore, it is crucial to ensure that public web applications are secure against such threats.

Developing a web application and developing a secure web application are two distinct tasks. While anyone can learn to create a web application, developing a secure web application requires specific knowledge and understanding of web security.

Research and interviews have indicated that deploying secure web applications is preferable to educating the public on the use of vulnerability and security scanner tools. However, even though vulnerability scanners are available, they are often expensive, making it difficult for small-scale companies to utilize them extensively. This financial barrier exacerbates the challenge of ensuring web application security. Consequently, there is a pressing need for affordable and effective tools that web application developers and quality assurance engineers can use to test for common vulnerabilities, such as SQL injections and Cross-Site Scripting, before deploying their final product for public use.

c. Project Objectives

Project objectives can be defined as goals that are specific, defined, and measurable that the project aims to achieve within a given time period. Project objectives help to guide the team in the right direction to success. These project objectives should be SMART [5].Which means they should be Specific, Measurable, achievable, relevant, and finally Time bound. These are some of the project objectives of the Vulnerability Scanner,

- To build a less complicated on-site scanning tool for web developers
- To detect SQL injection and XSS vulnerabilities
- To get feedback and questions from the users
- To develop a manageable tool with admin features
- To store user details and scan results along with the scanned URL in a database

II. LITERATURE REVIEW

The importance of websites and web applications in today's world was discussed in the document published by Kirti Sharma and Shobha Bhatt [6]. The paper covers what SQL injection attacks are and various types of SQL injection techniques with their intended attack. The document starts with a literature review based on questions from research, quality assessment and data samples. The author and the teams also discuss further techniques to prevent those mentioned various types of SQL injection attacks. Also, a tabular representation with grades of quality

evaluation. Lastly the team had also included research questions about SQL injection attacks along with their answers. This research paper gave an insight to understand the landscape of SQL injection attacks and strategies that can be used to prevent them. It also supports the need for vulnerability scanners to enhance the security of web applications.

A survey by Upasana Sharma and others provides a detailed analysis of existing detecting techniques for Cross-Site scripting attacks [7]. The paper explains what a Cross-Site scripting attack is by explaining how an attack happens and its consequences. Then the team categorize various detection methods of XSS attacks available in the literature for those attacks. Those detection methods are classified according to how they are deployed and further sub-categorized according to the analysis mechanism they employ. The survey also has documented pros and cons of each of those mentioned methods along with a list of scanning tools that can be used to detect Cross-site scripting attacks. This survey helps anyone who is interested to learn about XSS attacks because it is discussed in this paper what preconditions have to be met in order to launch a successful XSS attack. This knowledge is very useful to understand and design a detection mechanism to detect them.

The conference paper by Sabrina Tarannum and others explored web-based attacks and how they have grown in these past few years due to COVID-19 pandemic [8]. The paper discusses remote work and the importance of websites and web applications in day-to-day life. It is mentioned how web-based attacks have now become a significant part of digital forensics due to the growth of the number of attacks and how the attackers are gaining the capacity to go around the security measures and capable of causing distress for web application users. The paper shows that the biggest obstacle is to how to respond to unidentified threats. This paper is very useful as it talks about detection methods for various types of web-based attacks. Also helps to investigate and identify vulnerabilities and protection techniques along with potential future paths of online assault detection for research development.

Mr. Abhishek Kumar Baranwal had published a research paper on detecting SQL injection attacks and XSS attacks [9]. This document covers various kinds of SQL injection attacks and XSS attacks through a detailed survey and about what approaches should be taken to detect and prevent those attacks and assisted the development process of the scanner.

Shashank Gupta and B. B Gupta conducted a study on XSS attacks and had analyzed some major concerns for web applications and internet-based services such as health care, financial services, banking etc. [10] by referring to the Website Security Statistics Report of White Hat Security. In this paper, the taxonomy of XSS attacks was discussed and it also supplied a detailed analysis on exploitation, detection and prevention of XSS attacks.

Secubat is a web vulnerability scanner created by Stefan Kals, Engin Kirda, Christopher Kruegel and Nenad Jovanovic [11]. This scanner is similar to a port scanner and automatically analyzes websites for SQL Injections and XSS. To test Secubat's accuracy, they conducted scans on one hundred websites and found exploitable vulnerabilities on most of them. According to the document, more than fifty of those websites requested more information about their status of security.

This literature review was created according to the research that was conducted to get a clear idea on how to create a vulnerability scanner. The chapter was organized around several key themes such as a study on SQL injection and XSS attacks, a study on XSS attacks, why vulnerability scanners are needed, and an existing vulnerability scanner named "Secubat".

III. METHODOLOGY

d. Methodology

This chapter outlines the methodology used for developing the vulnerability scanner, aimed at assisting web developers and quality assurance engineers in identifying vulnerabilities such as SQL injection and Cross-Site Scripting (XSS) prior to deployment.

The development process follows the Software Development Life Cycle (SDLC), which includes Planning, Analyzing, Designing, and Implementation [12]. Adhering to the SDLC provides a structured approach to ensure the creation of a robust software solution. To enhance this process, the Agile methodology was adopted. Agile emphasizes iterative progress and flexibility, critical for adapting to evolving requirements and delivering high-quality software.

Agile methodology involves iterative cycles known as sprints, each comprising planning, design, development, testing, deployment, review, and launch. In the initiation phase, project objectives and requirements are defined, creating a backlog that guides subsequent work. Sprint planning involves selecting and breaking down backlog items into tasks for development.

Continuous testing throughout each sprint helps identify and resolve issues early, ensuring a more stable final product. The review phase at the end of each sprint provides opportunities for stakeholder feedback and iterative improvements.

By leveraging Agile methodology, the development of the vulnerability scanner benefits from regular refinement and alignment with project objectives, ultimately producing a tool that effectively meets security requirements and enhances web application safety.

e. System Overview

3.2.1) High Level Architecture

The below image is the high-level architectural representation of this on-site web vulnerability scanner. The user interacts with the scanner through the user interfaces which were developed using Angular, by providing feedback and URL and receiving feedback and results. HTTP requests are sent to the backend Flask server to pass the URL to the python scanner and conduct the scan. The results are then passed back to the user interface and MongoDB for storing, which also stores feedback and user details.

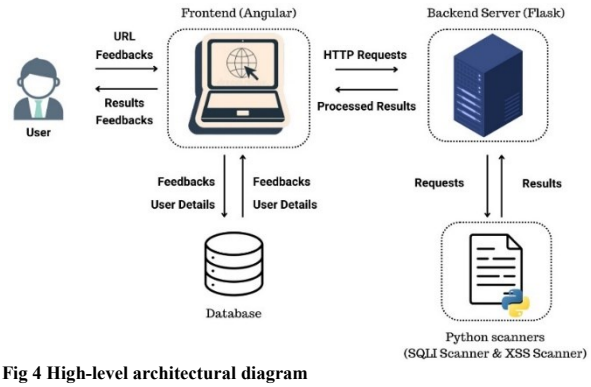


Fig 4 High-level architectural diagram

3.2.2) Scanning Process

When a registered user enters the URL they want to conduct the scan on, a Python function parses all the HTML components and returns all the forms that were found. Then the next function extracts details from every form found such as action URL, input fields and method. Then a payload is injected to all those identified input fields in the detected forms and will be submitted. If a payload is present in the response, the scanner will give the output as that URL is vulnerable along with the details such as in which form, input field, the vulnerability is present, the method etc.

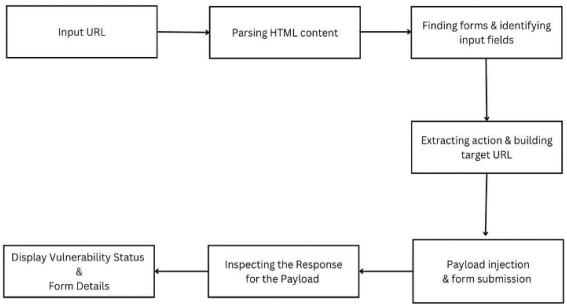


Fig. 2. Scanning Process

3.2.3) User Features

All the scans that were conducted through both XSS and SQL injection scanners are saved in a database along with the results of each scan. The admin can access these details and delete them if needed. Also, there are features for the admin such as “Users” to view user details and manage them and “FAQ” to view or delete FAQs one by one.

Both registered and unregistered users can send and view FAQ to the vulnerability scanner while only the registered users have the privilege to conduct scans through the scanner after logging in using their user credentials.

f. Implementation

Both SQL injection and cross-site scripting scanners were developed using the Python programming language. To scan for SQL injection and XSS vulnerabilities, first the HTML content of the URL should be parsed to identify the elements in it. BeautifulSoup module which is available as a python library was used to do the HTML parsing and then the forms are detected.

```
def get_all_forms(url):

    soup = bs(requests.get(url).content, "html.parser")
    return soup.find_all("form")
```

Fig 5 Parsing all HTML forms

After detecting all the forms in the URL, a function works to extract form details such as action, method and input fields. It is a very important step in this scanning process considering this information is critical when injecting payload and submitting it to get accurate scan results. Injecting and form submission is done by the next python function using the form details that were gathered by the previous function.

```
if form_details["method"] == "post":
    return requests.post(target_url, data=data)
else:
    return requests.get(target_url, params=data)
```

Fig. 4. Method selection for form submission with payload

Actual scanning is done by the next function where the response after submitting the form with malicious payload is monitored for any reflected payload. If there are no reflected payloads, the URL is not vulnerable but if it is, all the form details that were gathered during the scanning process such as action, method, input field, the form the vulnerability was present and the payload that triggered the vulnerability are supplied in the final scanner result.

The rest of the web application is developed using MEAN (MongoDB, Express, Angular and Node) stack and JWT tokens were used for secure authentication. Flask framework was used to connect the python scanners with the Angular user interfaces. When a user enters an URL into the input filed in the user interface, that URL will be passed to the python scanner through the Flask framework. The scanning process starts after the URL was received by the python scanner and when it is done, the result of the conducted scan is passed back to the interface for the user to see.

The scan result is then stored in MongoDB along with the URL and the user ID where the administrator can monitor and has the ability to delete them if wanted. Visual Studio Code was used as the code editor to develop this project.

IV. OBSERVATIONS AND DISCUSSIONS

The below images show the user interfaces where the users can enter the URL to initiate the scans for SQL injection and XSS vulnerabilities. After the users click scan button, the URL is passed to the python scanner to scan it. During the scanning process, “scanning…” is displayed to inform the users that the scan is proceeding. The scanning time will depend on the number of forms that were found within the URL because the payload will be injected to each, and every form found. The results appear below, stating whether the URL is vulnerable or not along with form details and the payload that triggered the vulnerability if the URL is in fact vulnerable.

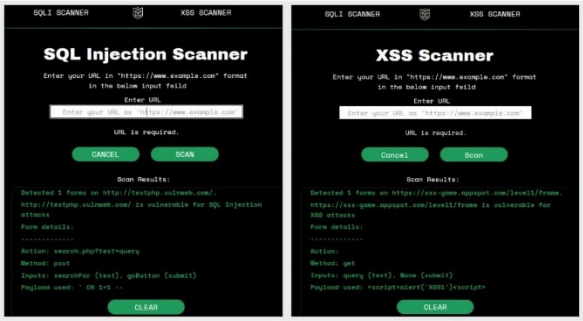


Fig. 5. *SQLI and XSS scan results*

Below image shows the admin interface where the admin can view all the scans that were conducted along with their result and the user ID. This approach has many advantages, such as making it easier to conduct audit trials, historical analysis and accountability. It allows you to track who initiated each scan and the subject that was tested, progress tracking etc. This function enhances user experience making the vulnerability scanning process more reliable.

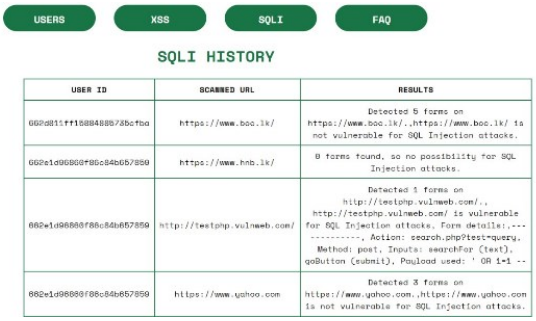


Fig. 6. *SQLI scan history - administrator*

The “Users” function provides the ability to delete users of the scanner one by one. This approch plays an important role in managing user accounts and maintaining overall security of the system. This allows the administrator precise control over who has access to the system and the ability to delete accounts of the users who no longer need it. This reduces the risk of unauthorized access, better tracking and accountibility. Through the “FAQ” function, the administrator can view feedback and questions from users and able to delete them one by one.

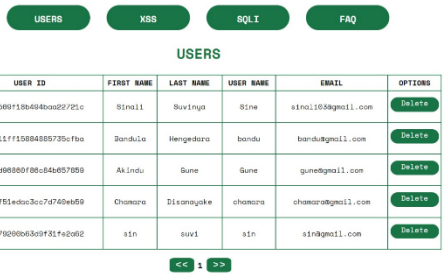


Fig 7 User details – administrator

V. CONCLUSION

In conclusion, this vulnerability scanner is a vital tool tailored specifically for small-scale software companies, designed to identify and address SQL injection and Cross Site Scripting (XSS) vulnerabilities before their web applications are launched. This scanner offers an accessible, user-friendly solution for small firms, enabling them to

perform crucial security testing affordably and efficiently. By focusing on detecting critical vulnerabilities, the tool helps ensure that applications are robust and secure against common threats. The inclusion of features such as user feedback mechanisms and administrative controls enhances its practicality and usability. Furthermore, the implementation of JSON Web Tokens (JWT) for authentication underscores the tool's commitment to security by protecting against Cross Site Request Forgery (CSRF) attacks. Overall, this on-site, simple vulnerability scanner contributes significantly to improving web application security for small-scale software companies, providing a reliable means of testing and fortifying products before their public release.

REFERENCES

[1] L. Baresi, F. Garzotto and . P. Paolini , "From Web Sites to Web Applications: New Issues for Conceptual Modeling," in Springer, Berlin, Heidelberg, 2001.

[2] V. Unic, "8 Types of Web Application Attacks and Protecting Your Organization," 10 May 2023. [Online]. Available: https://brightsec.com/blog/8-types-of-web-application-attacks-and-protecting-your-organization/.

[3] C. Disanayake, S. Dilhara and N. S. Dharmaratna, "Rationalized Security Frameworks for Web Applications," Homagama, 2023.

[4] "OWASP Top Ten," 2021. [Online]. Available: https://owasp.org/www-project-top-ten/.

[5] O. Ogbeiwi, "Why written objectives need to be really SMART," 2017.

[6] K. Sharma and S. Bhatt, "SQL injection attacks - a systematic review," International Journal of Information and Computer Security, vol. 11, 2019.

[7] U. Sarmah, D. K. Bhattacharyya and J. K. Kalita, "A survey of detection methods for XSS attacks," Journal of Network and Computer Applications, vol. 118, pp. 113-143, 2018.

[8] S. Tarannum, S. M. M. Hossain and . T. Sayeed , "Cyber Security Issues: Web Attack Investigation," in Springer, Cham, 2023.

[9] A. K. Baranwal, "Approaches to detect SQL injection and XSS in web," 2012.

[10] S. Gupta and B. B. Gupta, "Cross-Site Scripting (XSS) attacks and defense mechanisms: classification and state-of-the-art," International Journal of System Assurance Engineering and Management, vol. 8, p. 512–530, 2015.

[11] S. Kals, E. Kirda, C. Kruegel and N. Jovanovic, "SecuBat: a web vulnerability scanner," pp. 247 - 256, 2006.

[12] E. A.Altameem, "Impact of Agile Methodology on Software Development," 2015.

Enhancing Security with Biometric and Heart Rate Data Integration: A Comprehensive Review

Gayan Perera
Department of Computer & Data Science
NSBM Green University Town
Homagama, Sri Lanka
gayanp@nsbm.ac.lk

Mandira De Silva
BSc(Honors) Software Engineering (UG)
NSBM Green University Town
Homagama, Sri Lanka
dumdsilva@students.nsbm.ac.lk

Shwetha Mandakini
BSc(Honors) Software Engineering (UG)
NSBM Green University Town
Homagama, Sri Lanka
pgsmpremakumara@students.nsbm.ac.lk

Kavindi Kinkini
BSc(Honors) Software Engineering (UG)
NSBM Green University Town
Homagama, Sri Lanka
wakkinkini@students.nsbm.ac.lk

Tracey Johnson
BSc(Honors) Software Engineering (UG)
NSBM Green University Town
Homagama, Sri Lanka
tsjpaul@student.nsbm.ac.lk

Nethmini Navodya
BSc(Honors) Software Engineering (UG)
NSBM Green University Town
Homagama, Sri Lanka
kgnnjayamani@students.nsbm.ac.lk

***Abstract*—Fingerprint biometrics and heart rate biometrics are foundational technologies in personal identification and authentication, each offering distinct advantages in biometric security. This review provides a comprehensive analysis of both technologies, drawing on current research to evaluate their respective strengths and limitations. Fingerprint biometrics, known for their high uniqueness and reliability, are widely used in various applications, from securing personal devices to large-scale government identification programs. However, they are not without vulnerabilities, particularly their susceptibility to spoofing and privacy concerns. On the other hand, heart rate biometrics, derived from electrocardiogram (ECG) signals, present a novel approach to authentication, offering continuous verification and a robust defense against spoofing. Despite their potential, heart rate biometrics face challenges such as accuracy issues due to physiological variations and significant privacy concerns. This paper aims to identify the gaps in current biometric security measures and propose innovative solutions for enhancing data center security by integrating heart rate data into existing fingerprint biometric systems. The proposed multi-factor authentication approach promises to enhance security by leveraging the strengths of both modalities while addressing their respective weaknesses.**

Keywords—Biometrics, Heart Rate Data, Fingerprint Recognition, Security, Continuous Authentication

I. INTRODUCTION

Biometric technology is revolutionizing identity authentication by leveraging unique biological traits to provide secure and reliable identification methods. Among various biometric modalities, fingerprint biometrics has emerged as one of the most extensively used and researched technologies. Their high degree of reliability and ease of use have made them indispensable in a wide range of applications, from personal device security to large-scale government identification programs[1]. The distinctiveness of fingerprint patterns, even among identical twins, ensures their effectiveness in providing accurate and reliable identification throughout an individual's lifetime [2].

Despite their widespread adoption and advantages, fingerprint biometrics are not without their limitations. The susceptibility of fingerprint systems to spoofing and forgery is a significant security concern, as artificial fingerprints can deceive even advanced recognition systems [3]. Additionally, the quality of fingerprints can be affected by various factors such as age, occupation, and environmental conditions, leading to recognition difficulties and impacting the overall effectiveness of the system[4]. Privacy concerns also arise from the collection and storage of fingerprint data, as compromised fingerprints cannot be easily changed, posing severe risks in the event of data breaches [5].

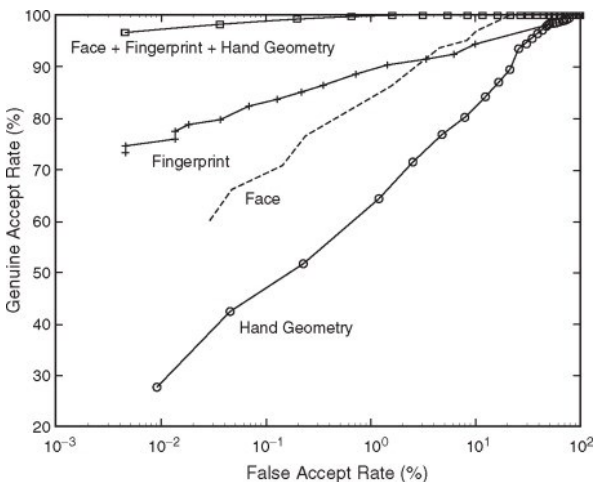


Fig. 6: Multimodal biometric

[22]

In contrast, heart rate biometrics offer a novel approach to personal identification and authentication. Derived from electrocardiogram (ECG) signals, heart rate biometrics provide continuous authentication by leveraging the inherent and continuous presence of heartbeats. This modality is less susceptible to external manipulations compared to other biometric systems, as replicating an individual's unique heartbeat is significantly more challenging [6]. The integration of heart rate biometrics into security systems can enhance security by providing real-time verification and continuous monitoring, ensuring that the authenticated user remains present [7].

However, heart rate biometrics also face inherent challenges. Variations in heart rate patterns due to medical conditions or physical activity can lead to false negatives or positives, affecting the reliability of these systems [8]. The sensitive nature of heart rate data raises significant privacy concerns, as its misuse can have severe consequences[9]. Additionally, integrating heart rate biometrics into existing security frameworks requires navigating complex regulatory and ethical landscapes, with issues related to data ownership and consent necessitating robust legal frameworks[10].

Fig. 1 clearly illustrates that the collection of more biometric data towards a multimodal biometric system can significantly enhance the genuine acceptance rate compared to using a single biometric data point of an individual.

This review aims to explore the potential of combining fingerprint and heart rate biometrics to enhance security measures, particularly in data centers. By addressing existing gaps and proposing innovative solutions, this paper seeks to provide a comprehensive understanding of the strengths and limitations of both biometric modalities and their integration to develop a more robust and reliable multi-factor authentication system.

II. RESEARCH OBJECTIVES

f. Investigating the Security Benefits of Integrating Heart Rate Data into Existing Biometric Systems

i. Evaluate how heart rate data can enhance the security of existing fingerprint biometric systems by preventing biometric spoofing attempts.

ii. Explore the potential of using heart rate data for stress detection to identify compromised users or attempted security breaches, thereby providing proactive security measures.

g. Assessing the Feasibility of Heart Rate Response-Based Continuous Low-Impact Challenges

i. Determine the feasibility of maintaining secure access through continuous low-impact challenges based on heart rate responses without significantly impacting user movement or experience.

ii. Address the reliability and predictability concerns associated with heart rate response-based challenges, ensuring the system's effectiveness across diverse user populations.

h. Developing a Multi-Factor Authentication System

i. Design and propose a multi-factor authentication system that combines fingerprint and heart rate biometrics to leverage the strengths of both modalities while mitigating their respective weaknesses.

ii. Implement robust data encryption protocols and diverse training datasets to enhance the security and accuracy of the combined biometric system.

III. LITERATURE REVIEW

Fingerprint biometrics has long been regarded as one of the most reliable and widely adopted methods of personal identification. The uniqueness of fingerprints, even among identical twins, ensures a high degree of reliability for identification purposes [1]. This inherent distinctiveness is further bolstered by the stability of fingerprint patterns throughout an individual's lifetime, making it a dependable biometric modality[2]. The ease of use and convenience associated with fingerprint recognition systems have contributed to their widespread adoption across various applications, from smartphone security to large-scale government identification programs [11]. The non-intrusive nature of fingerprint scanning enhances user acceptance, particularly when compared to more intrusive biometric methods such as iris or retina scans[12].

Technological advancements have significantly reduced the cost of fingerprint sensors, making them a

viable option for widespread adoption in consumer electronics and security devices [13]. Additionally, fingerprint recognition systems are known for their rapid processing and verification capabilities, which are crucial in high-throughput environments such as airports and critical infrastructure facilities [14]. This swift processing speed not only facilitates efficient access control but also helps deter unauthorized access attempts by minimizing wait times and bottlenecks at security checkpoints.

Despite these advantages, fingerprint biometrics are not without their limitations. One of the most significant vulnerabilities of fingerprint systems is their susceptibility to spoofing and forgery. Research has demonstrated that artificial fingerprints, created from materials such as silicone or gelatin, can successfully deceive fingerprint recognition systems, raising substantial security concerns [3]. Furthermore, the quality of fingerprints can be affected by various factors, including age, occupation, and skin conditions. For instance, manual laborers or elderly individuals may have worn or damaged fingerprints, leading to recognition difficulties [4]. Environmental conditions, such as moisture, dirt, and temperature, can also impact the effectiveness of fingerprint sensors [12].

The collection and storage of biometric data, including fingerprints, pose significant privacy concerns. Unauthorized access to or breaches of fingerprint databases could result in identity theft and misuse. Unlike passwords, fingerprints cannot be easily changed if compromised, making the consequences of data breaches particularly severe [5]. Integrating fingerprint biometrics into existing systems can be a complex and expensive undertaking, especially for organizations with outdated infrastructure. Ensuring interoperability between different fingerprint recognition systems and standards can pose additional challenges[15]. Moreover, fingerprint systems are not immune to errors, as they can generate false rejections (denying access to legitimate users) and false acceptances (granting access to unauthorized users) [15].

In contrast to fingerprint biometrics, heart rate biometrics offer a novel approach to personal identification and authentication. The inherent and continuous presence of heartbeats makes heart rate biometrics universally applicable and reliable for continuous authentication [16]. Heart rate biometrics are less susceptible to external manipulations compared to facial or voice recognition systems, as replicating an individual's unique heartbeat is significantly more challenging [6]. This internal nature of heart rate biometrics provides a robust defense against common spoofing tactics, enhancing overall security [17].

One of the most significant advantages of heart rate biometrics is their ability to facilitate continuous authentication. Unlike static methods such as passwords or fingerprint scans, which verify identity only at the point of entry, heart rate biometrics provide real-time credentials, ensuring that the authenticated user remains present. This continuous authentication is particularly beneficial for sensitive operations requiring ongoing verification of user presence [7]. The miniaturization of ECG technology, as seen in wearable devices like the Apple Watch, highlights the practicality of heart rate biometrics in everyday use. These portable devices can continuously monitor and authenticate users, blending seamlessly into daily activities while maintaining high security levels[18].

However, heart rate biometrics also come with inherent risks and challenges. The storage and transmission of biometric data, including heart rate data, pose significant security risks. If compromised, this data can be used for malicious purposes, as it cannot be altered like passwords [9]. Additionally, variations in heartbeat patterns due to medical conditions or physical activity could lead to false negatives or positives, impacting the reliability of heart rate biometric systems [8]. The integration of biometric data into various systems raises ethical and regulatory challenges, with issues related to data ownership, consent, and potential misuse necessitating robust legal frameworks [10].

When comparing fingerprint and heart rate biometrics, it becomes evident that each modality has its strengths and weaknesses. Fingerprint biometrics are well-established and technologically mature, but they exhibit limitations in terms of accuracy and privacy concerns. On the other hand, heart rate pulse recognition offers promising advancements in liveness detection and continuous verification but faces challenges in accuracy and privacy.

To harness the advantages of both biometric modalities while addressing their respective risks, several mitigation strategies can be employed. Implementing multi-factor authentication by combining heart rate and fingerprint biometrics can enhance security by mitigating the risks associated with each modality. Robust data encryption protocols for storing and transmitting biometric data are essential to reduce the risk of unauthorized access and misuse [19]. Training biometric systems on diverse datasets can minimize biases and improve.

IV. GAP IDENTIFICATION

Despite the advancements and widespread adoption of fingerprint biometrics, significant gaps remain that hinder their effectiveness. One of the primary vulnerabilities is their susceptibility to spoofing, even with advancements in liveness detection. This vulnerability raises substantial security concerns, particularly in high-security applications [3]. Additionally, the integration of fingerprint biometrics into existing systems poses challenges, especially for organizations with outdated infrastructure. Ensuring interoperability between different fingerprint systems adds to the complexity and cost of implementation [15]. Privacy concerns also persist, as compromised fingerprints cannot be

changed, leading to severe consequences in the event of data breaches [5].

Similarly, heart rate biometrics, while offering innovative solutions for continuous authentication and enhanced security, are not without their challenges. Variations in heart rate patterns due to medical conditions or physical activity can result in false negatives or positives, impacting the reliability of these systems [8]. The sensitive nature of heart rate data raises significant privacy concerns, as its misuse can lead to severe consequences [9]. Furthermore, integrating heart rate biometrics into existing security frameworks requires navigating complex regulatory and ethical landscapes, with issues related to data ownership and consent necessitating robust legal frameworks [10].

V. PROPOSED ENHANCEMENTS

To address these gaps and enhance biometric security, integrating heart rate data into existing fingerprint biometric systems presents a promising solution. This integration can provide additional layers of security by detecting stress or anxiety, which may indicate compromised users or attempted breaches. Real-time heart rate monitoring can offer continuous authentication, ensuring secure access without impacting user movement.

One innovative approach to enhancing security is the development of a multi-factor authentication system that combines fingerprint and heart rate biometrics. By leveraging the strengths of both modalities, this system can mitigate the weaknesses of each and provide a more robust and reliable authentication mechanism. For instance, while fingerprint biometrics offer high uniqueness and ease of use, their susceptibility to spoofing can be addressed by incorporating heart rate data, which is difficult to replicate. Similarly, the continuous authentication capability of heart rate biometrics can complement the rapid processing and verification speed of fingerprint systems, ensuring ongoing user verification and enhancing overall security.

To implement this multi-factor authentication system, robust data encryption protocols must be established to protect the storage and transmission of biometric data. Ensuring that data is anonymized and encrypted can reduce the risk of unauthorized access and misuse [19]. Additionally, training biometric systems on diverse datasets can minimize biases and improve accuracy across different demographics, further enhancing the reliability and security of the system [20]. Adhering to data protection regulations and ethical standards is crucial for maintaining trust and security. Organizations must implement transparent data handling practices and ensure users' consent is obtained and respected, addressing potential regulatory and ethical challenges [21].

In conclusion, the integration of heart rate data into existing fingerprint biometric systems offers a promising solution to enhance security in data centers and other high-security environments. By combining the strengths of both biometric modalities, a more robust and reliable multi-factor authentication system can be developed, addressing the limitations of each and enhancing overall security. Future research should focus on addressing the accuracy and privacy concerns associated with heart rate biometrics and developing comprehensive legal frameworks to ensure their ethical use.

PROPOSED ENHANCEMENTS

The following table presents proposed solutions aimed at mitigating the identified problems.

TABLE I: Proposed Solutions

Index	Stream	Problem	Suggested Solution
A	Investigating the security benefits of integrating heart rate data into existing biometric systems	How can heart rate data enhance the security of existing systems by preventing biometric spoofing attempts?	Integrating heart rate data to create a dual-authentication mechanism that requires a live,

			human heartbeat
		The potential of using heart rate data for stress detection to identify compromised users or attempted security breaches is challenged by the difficulty in accurately detecting stress levels with biometric data and the risk of false positives	Develop algorithms to analyze heart rate variability (HRV) and correlate it with stress levels. Use machine learning models trained on diverse datasets to improve detection accuracy and reduce false positives
B	Accessing the feasibility of heart rate response-based continuous low-impact challenges	The feasibility of maintaining secure access through continuous, low-impact challenges based on heart rate responses is complicated by the need	Design low-impact heart rate monitoring devices that seamlessly integrate with user activity. Use non-intrusive sensors and ensure data collection is passive and does not interrupt user activities
		to do so without significantly affecting user movement or experience	
		The reliability and predictability of heart-rate response-based challenges raise concerns, particularly regarding the system's effectiveness across diverse user populations	Conduct extensive testing across diverse demographics to account for variability. Implement adaptive algorithms that can learn and adjust to individual user patterns, improving reliability and predictability
C	Developing a multi-factor authentication	Design and propose a multi-factor authentication	Develop a user-friendly interface

	n system	on system that integrates fingerprint and heart rate biometrics to capitalize on the strengths of each modality while mitigating their respective weakness	that allows easy and quick authentication using both modalities. Optimize the system architecture to ensure minimal delay and high efficiency during the authentication process.
		Implement robust data encryption protocols and utilize diverse training datasets to enhance the security and accuracy of the combined biometric system, ensuring reliable performance and protecting sensitive information	Use advanced encryption protocols to secure data during transmission and storage. Collect and utilize diverse datasets for training machine learning models to enhance the accuracy
			and robustness of the biometric system

CONCLUSION

This review highlights the vulnerabilities of fingerprint biometrics to spoofing and proposes the integration of heart rate data as an innovative solution to enhance security. While heart rate biometrics offer significant advantages in continuous authentication and security, they also pose privacy and regulatory challenges. The proposed multi-factor authentication approach, combining fingerprint and heart rate biometrics, promises to enhance security in data centers by providing a more robust and reliable authentication mechanism. Future research should focus on addressing the accuracy and privacy concerns associated with heart rate biometrics and developing comprehensive legal frameworks to ensure their ethical use.

REFERENCES

[1] D. Maltoni, D. Maio, A. K. Jain, and S. Prabhakar, "Handbook of Fingerprint Recognition," Springer.

[2] A. K. Jain and J. Feng, "Latent fingerprint matching," *IEEE Trans Pattern Anal Mach Intell*, vol. 33, no. 1, pp. 88–100, 2011, doi: 10.1109/TPAMI.2010.59.

[3] U. Uludag and A. K. Jain, "Attacks on Biometric Systems: A Case Study in Fingerprints." [Online]. Available: http://biometrics.cse.msu.edu

[4] J. Ashbourn, "Practical Biometrics: From Aspiration to Implementation," Springer.

[5] A. Cavoukian and A. Stoianov, "Biometric Encryption: A Positive-Sum Technology that Achieves Strong Authentication, Security AND Privacy ABOUT THE AUTHORS," 2007. [Online]. Available: www.ipc.on.ca

[6] D. Jones and H. Williams, "Overcoming the challenges of biometric fraud," *Journal of Digital Security, 14(3), 177-190*, 2021.

[7] A. , et al. Miller, "Continuous authentication with heartbeat biometrics," *Journal of Biometric Science, 10(1), 56-72*, 2021.

[8] M. Clark and P. Rodriguez, "Accuracy and bias in biometric systems," *Journal of Biometric Research, 8(2), 145-159*, 2021.

[9] G. Taylor and K. Smith, "Privacy issues in biometric authentication," *Journal of Information Privacy, 7(3), 112-128*, 2019.

[10] E. Wilson and M. Green, "Legal frameworks for biometric data," *Journal of Law and Technology, 13(4), 321-338*, 2020.

[11] J. Galbally, S. Marcel, and J. Fierrez, "Biometric antispoofing methods: A survey in face recognition," *IEEE Access*, vol. 2, pp. 1530–1552, 2014, doi: 10.1109/ACCESS.2014.2381273.

[12] T. Fronczak and J. Pritchard, "Biometrics for Security and Privacy," IEEE Security & Privacy.

[13] A. Kumar and D. Zhang, "Hand-Geometry Recognition Using Entropy-Based Discretization," IEEE Transactions on Information Forensics and Security.

[14] A. K. Jain, K. Nandakumar, and A. Nagar, " Fingerprint matching and identification systems: components and algorithms," *Pattern Recognit*, 2000.

[15] N. K. Ratha, J. H. Connell, and R. M. Bolle, "Enhancing Security and Privacy in Biometrics-Based Authentication Systems," *IBM Systems Journal*, 2001.

[16] J. , et al. Smith, "The universality of heartbeat biometrics," *Human Factors in Biometric Security, 16(1), 45-58*, 2022.

[17] A. Brown and K. Davis, "Enhancing security with heartbeat biometrics," *Cybersecurity Journal, 15(3), 98-112*, 2020.

[18] L. Johnson and S. Lee, "Portability of biometric devices in everyday use," *Wearable Technology Journal, 9(1), 34-47*, 2022.

[19] R. Chen and L. Zhao, "Data encryption in biometric systems," *International Journal of Information Security, 12(4), 213-227*, 2021.

[20] P. Nguyen and R. Martin, "Reducing bias in biometric systems," *Journal of Machine Learning and Data Mining, 15(2), 201-217*, 2020.

[21] S. Adams and J. Evans, "Ethical considerations in biometric data management," *Journal of Data Privacy, 10(2), 233-245*, 2019.

[22] Shaveta Dargan and Munish Kumar, "A comprehensive survey on the biometric recognition systems based on physiological and behavioral modalities," Science Direct. [Online]. Available: https://www.sciencedirect.com/topics/computer-science/multimodal-biometric-system

Source Code Vulnerability Detection Using Static Analysis and Pattern Matching

Jithmal Danusha Pitiyegedara
Department of Network and Security,
NSBM Green University,
Homagama, Sri Lanka
pgjdanusha@students.nsbm.ac.lk

Isuru Sri Bandara Nishshanka
NSBM Green University,
Homagama, Sri Lanka.
isuru.s@nsbm.ac.lk

***Abstract*- In the rapidly evolving field of web development, ensuring the security of web applications is paramount. Most of the vulnerabilities occur due to lack of best practices during coding. Traditional security testing methods, which are predominantly reactive, tend to identify vulnerabilities in source code only after substantial portions of the application have been developed. This reactive approach significantly increases the complexity and cost of mitigating security risks. This research addresses this challenge by integrating security directly into the development process through a VS(Visual Studio) Code Extension by performing a real-time Static Analysis of source code to detect vulnerabilities which are written using JavaScript language. The tool can identify common security vulnerabilities such as SQL Injection (SQLi), Cross-Site Scripting (XSS), Cross-Site Request Forgery (CSRF) using the pattern matching techniques of static analysis. By integrating the tool directly into VS Code, it can provide developers with immediate feedback on security risks. The tool is not only useful for identifying vulnerabilities but also enhances the developer's overall coding practices. The integration of the tool into the development environment will reduce time and resources spent on post- development security testing. In conclusion, this tool represents a significant step forward in the effort to embed security within the software development lifecycle. Its development and deployment will underscore the effectiveness of proactive security practices, offering a model that could be replicated and extended to other development environments and programming languages.**

Keywords—static analysis, pattern matching, JavaScript, source code, vulnerabilities, proactive testing, vs code

I. INTRODUCTION

In the landscape of cyber technology, security has become a paramount concern, especially with the increasing complexity and functionality of applications. As computer applications evolve, they often incorporate various services that handle sensitive user data or perform critical operations, such as payment processing, personal data management, and other interactive functionalities. These enhancements, while improving user experience and functionality, also broaden the attack surface, making applications susceptible to a variety of threats.

Early threats like viruses and worms have given way to more complex attacks such as phishing, ransomware, and advanced persistent threats (APTs). The proliferation of Internet of Things (IoT) devices and the widespread adoption of cloud computing have introduced new vulnerabilities, making these systems prime targets for cybercriminals. Modern threats now leverage advanced techniques like AI and machine learning to evade detection and automate attacks, posing significant risks to both individuals and organizations. As the digital landscape continues to expand, the importance of robust cybersecurity measures is more critical than ever.

Cyber-Related Attack	Year	Attack Spread Method	Consequences
Vladimir Levin's Attack the Citibank1	1994–1995	unknown	around 10 million dollars were stolen
Melissa Virus	1999	used users' trust to click an email attachment	billions of dollars were lost in many countries
ILOVEYOU Worm3	2000	used users' trust to click an email attachment	more than 45 million computers were infected
MyDoom worm4	2004	used attention-grabbing subjects by email, such as errors, tests, etc.	DDoS attacks by allowing remote access was launched
Zeus Trojan	2007	spam email with drive-by downloads	login details such as email and bank accounts were stolen
Stuxnet Worm	2010	attack the programmable logic unit by stealing source codes	control of industrial processes was taken
Attack on USA Natural Gas Pipeline	2012	accessing confidential information through phishing	security credentials were stolen
Mirai Malware	2016	vulnerability of IoT devices was exploited	DDoS attacks were launched
WannaCry Ransomware	2017	windows vulnerability was exploited	computer hard drives were encrypted, and 150 countries were affected
Emotet Trojan	2018	emails in the form of spam and phishing campaigns	sensitive information such as credit card details was stolen
MyFitnessPal	2018	software vulnerability was exploited	150 million users were affected
Ransomware Attack on Magellan	2020	emails in the form of spam and phishing campaigns	health data of 365,000 patients were stolen
CovidLock Ransomware	2020	exploited users' trust by using COVID-19 statistic	android devices' data were encrypted, and data access was denied
Accellion Supply Chain Attack	2021	third-party vulnerabilities were exploited	confidential data from large organizations were stolen
Kaseya Ransomware Attack	2021	zero-day exploits were used	around 1500 companies' data were compromised with the request of 50,000 to 5 million dollar ransoms per victim

Fig. 1. Significant risks on AI

When looking at the Fig.1. [1] and other related internet sources, it can conclude that most of the attacks are related to email and web security. The history of web security threats dates to the early days of the internet when simple HTML websites transitioned into complex web applications. Initially, security was not a primary concern, but as web applications began to handle more sensitive information and perform critical functions, the importance of securing these applications became evident.

Numerous security tools have been developed to address the growing web security threats. These tools are generally categorized into Static Application Security Testing (SAST) and Dynamic Application Security Testing (DAST) methodologies[4]. While these tools are effective, they predominantly focus on the post-development stages of applications. Their reactive nature means vulnerabilities are often identified only after they have been introduced into the codebase, leading to higher remediation costs and potential delays in deployment.

Given the limitations in existing security tools, there is a clear and present need for a proactive approach that integrates security directly into the development phase of web applications. This research emerges as a solution to this need, providing a mechanism using Static Analysis to detect and mitigate vulnerabilities during the coding process itself. This proactive approach ensures that vulnerabilities can be identified and rectified in real-time, significantly reducing the chances of security risks in the later stages of development.

II. Literature Review

The proliferation of web-based applications has necessitated the development of advanced security tools to protect against a wide range of vulnerabilities. As the complexity of web applications increases, so does the variety of attack vectors. In response, security tools and methodologies have evolved to detect and mitigate potential threats. Two main types of tools Static Application Security Testing (SAST) and Dynamic Application Security Testing (DAST) have become foundational in the field of application security[4].

A. Static Application Security Testing (SAST)

SAST tools analyze source code to detect vulnerabilities without executing the program. These tools scan the entire codebase, including third-party libraries and dependencies, to identify patterns indicative of security flaws[5].

1) Advantages

SAST tools can identify vulnerabilities early in the development phase, allowing developers to address issues before deployment.
These tools provide a thorough analysis of the codebase, enhancing the security posture of the application.
SAST tools can be integrated into CI/CD pipelines, facilitating automated security checks during the development process.

2) Limitations:

These tools may generate false positives, leading to unnecessary remediation efforts.
While effective at code analysis, SAST tools might not detect issues related to runtime behavior or server-side configurations.

B. Dynamic Application Security Testing (DAST)

DAST tools interact with a running application, simulating real-world attacks to identify vulnerabilities. This approach involves testing the application's responses to various attack vectors, providing an assessment of its security in an operational state[5].

1) Advantages:

DAST tools simulate actual attack scenarios, providing a realistic evaluation of the application's security posture.
These tools assess the application's runtime behavior, configurations, and interactions with external systems.

2) Limitations:

DAST tools do not have access to the source code, which can restrict their ability to pinpoint specific vulnerabilities.
These tools are typically used later in the development lifecycle, which can make the remediation process more difficult and costly.

According to previous research done by researchers by testing both SAST and DAST tools, SAST tools found more Vulnerabilities when compared to DAST tools. The below Figure 2 represents the result outputs[5].

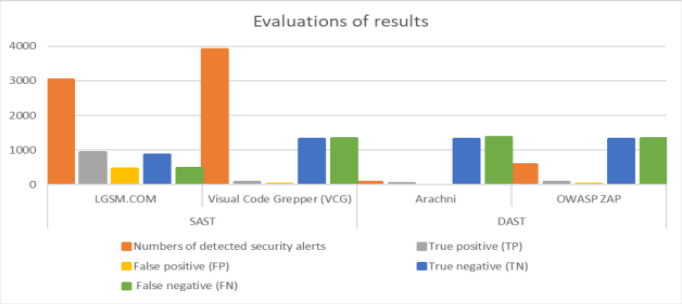


Fig. 2. Results Outputs

The figures show that existing DAST tools are highly effective in scanning the source code for vulnerabilities, but they typically operate post-development or during the later stages of the development lifecycle. This often results in vulnerabilities being detected only after significant portions of the application have been developed, which can lead to costly and time-consuming remediation processes. Moreover, these tools generally serve quality assurance teams rather than developers directly involved in the coding process. So, for the requirement addressed earlier to detect and mitigate code vulnerabilities during development phase, the SAST approach is most suitable.

II. Static Code Analysis

Static analysis refers to the examination of code without executing it, aiming to find potential vulnerabilities such as SQL Injection (SQLi), Cross-Site Scripting (XSS), and unsafe variable declarations[6]. Let's see how it leverages

static analysis to scrutinize JavaScript code for known vulnerability patterns.

A. Regex for Pattern Matching

Regex (Regular Expressions) is a powerful way for identifying specific patterns within text. In the context of the research, regex patterns are employed to match code snippets that correspond to known vulnerability patterns written in JavaScript[7]. For the explanation, let's take user input sanitization during development. Consider a JavaScript variable declaration like below,

var userInput = getUserInput();

Static analysis inspects the code to detect if 'userInput' is utilized in a manner that could lead to security issues. For instance, if 'userInput' is directly incorporated into an SQL query without sanitization, it could be susceptible to SQL injection attacks[8]. The vulnerability triggers when the un sanitized 'userInput' variable is concatenated in to a SQL query like below,

*var query = "SELECT * FROM users WHERE username = '"*

+ userInput + "';

Now because of the concatenation of 'userInput' directly without sanitizing is vulnerable, using a regex pattern, it's capable of performing a pattern matching to detect the patterns with concatenated user input in a SQL query. Next, to identify a SQL pattern like above we can implement a regex pattern. For example, a regex to detect unsanitized variable usage in an SQL query might look like this:

/SELECT\s+\s+FROM\s+\w+\s+WHERE\s+\w+\s*=\s*['"]?\s*\w+\s*['"]?/i*

This pattern leverages the power of regular expressions to match common SQL query structures that involve potentially unsafe variable concatenation. The characters which is included in the regex patterns are known as Metacharacters which are differ from language to language. Other than the metacharacters the regex can also include the special string literals like 'SELECT', 'FROM', 'WHERE' to identify the code snippet. Table 1 [9] represents the further breakdown of the above regex patten,

TABLE 1. REGEX BREAKDOWN

'["']?'	Matches zero or one single or double quote character.
'/i'	The i at the end indicates case-insensitive matching.

For brief analysis refer the book *'Regular Expression Pocket Reference by Tony Stubblebine'* which includes basic Metacharacters which used to create Regex Content which is supported for JavaScript[9]. To learn more further information about how regex work and other Metacharacters it can be found from online sources like regex101, w3schools etc.

Class	Meaning
`[...]`	A single character listed, or contained within a listed range.
`[^...]`	A single character not listed, and not contained within a listed range.
`.`	Any character except a line terminator, `[^\x0A\x0D\u2028\u2029]`.
`\w`	Word character, `[a-zA-Z0-9_]`.
`\W`	Nonword character, `[^a-zA-Z0-9_]`.
`\d`	Digit character, `[0-9]`.
`\D`	Nondigit character, `[^0-9]`.
`\s`	Whitespace character.
`\S`	Nonwhitespace character.

Table 40. JavaScript anchors and other zero-width tests

Sequence	Meaning
`^`	Start of string, or the point after any newline if in multiline match mode, `/m`.
`$`	End of search string, or the point before a string-ending newline, or before any newline if in multiline match mode, `/m`.
`\b`	Word boundary.
`\B`	Not-word-boundary.
`(?=...)`	Positive lookahead.
`(?!...)`	Negative lookahead.

Fig. 3. JavaScript Anchors

By referring to these sources there is a good possibility to make a complete algorithm to detect different kinds of vulnerable code patterns including SQLI, XSS, CSRF and other related vulnerabilities. The algorithm will include all possible vulnerable code patterns which are likely to be bad coding practices done by developers. After programming the algorithm, it can be implemented into a tool, software, third party application or to a programming IDE(vs code,

sublime text etc.) which could be most appropriate because the development of web systems is done by using IDE s.

III. LOGIC/ALGORITHM

Creating a complete and precise algorithm and integrating it to a working tool includes several steps. As discussed above, including regex for pattern matching it has a great potential for making a complex algorithm which covers the mentioned vulnerabilities earlier. The complex algorithm is consist of two main sections. As the vulnerability in a written code always comes with a variable declaration, the variables declared in a code space is important to be considered in the algorithm. Other than variable declarations the next consideration is likely to be function declarations. These two types of declarations are the most to be the starting point of known vulnerabilities. So identifying and tracking the declared variables and functions mainly using the algorithm will covers lots of vulnerabilities. Lets see the step-by step approach.

A. *Variable Tracking*

The first step would be identifying the variable declarations inside the written code space. When considering with JavaScript, using regex the algorithm can find declarations for like 'var', 'let' and 'const' keywords which are the ways for declaring variables. Then using the variable name it can be tracked to see if the raw variable is used somewhere which is leads to a vulnerability. For example consider the below code sample.

```
const Input1 = userInput1;
document.getElementById('output').innerHTML
= Input1;
```

In here, a user variable name 'Input1' has been declared which the assigned value for 'userInput1' is unknown. Then the input is directly assigned to a *'innerHTML'* inbuilt element which is vulnerable in web development[10]. For example if the unknown userInput1 is something like in below, then the code automatically opens an alert box in the browser window saying '*XSS*' when the *document.getElementById()* runs by the compiler which we can referred to as a XSS attack.

```
const userInput1 =
"<script>alert('XSS')</script>";
```

So to mitigate the kinds of attacks that occurs due to these kind of variable declarations it's important to sanitize inputs before using them directly. To detect these kind of unsafe code patterns the variable tracking algorithm plays a major role.

B. *Function Tracking*

Same as variable tracking, function tracking is also important for detecting vulnerable code patterns. Using regex, the algorithm will detect function declarations and using function tracking it can be analyzed for unsafe use of the function in the code space. Using the variable and function tracking algorithms it is possible to cover lots of vulnerable code patterns which is related to SQLi and XSS during web development.

Apart from variable and function tracking, we can also program other algorithms for detecting CSRF vulnerabilities which initiates with API declarations, e-mail related vulnerabilities which related to email header declarations[11], and many other types.

III. IMPLEMENTATION

With the complete algorithm which covers the identified vulnerabilities according to the research, to make the idea practical it can be implemented into a software program to analyze a code written by using a programming language(in this case JavaScript) to test the functionality of the concept. With the idea of technology behind the research as static analysis the tool should be capable of running real-time in a code space environment. When considering code space, the programming IDE s plays a major role in development stage. From the information gathered with surveys and other requirement analysis methods among university students, we asked what IDE is used by most students for their development for applications. The data shows like in Figure 4.

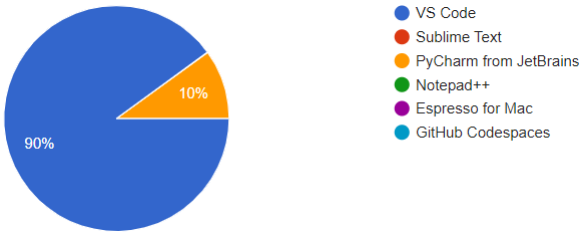


Fig. 4. Used IDE applications

It's clear that not only in our surveys the internet sources are also saying that the VS code is the most used IDE. So implementing the tool which

supports VS code as the initial target will be more suitable to test the tool. The VS code offers its extension API which developers can use to implement different extensions to run on VS code IDE. This is a perfect opportunity to implement the static code analysis algorithm which programed into an extension to run it real time on the VS code IDE. This will cover lots of userbase which uses the VS code and it can be use to successfully test and improve the quality of the algorithm as well.

Other than this approach the algorithm can also be implemented with a web based approach or any other kind of preferred way.

IV. TESTING AND CONCLUSION

Testing and validation are critical steps in the development process to ensure that the tool effectively identifies and helps mitigate security vulnerabilities. This chapter presents the results of testing the tool with various code snippets that contain common web security vulnerabilities. The functionality of the system was verified through these tests, demonstrating its capability to detect and warn about insecure code. The testing phase demonstrated that the system performs its intended functions effectively. The core functionality of the tool is to provide real-time analysis of code and detect potential security vulnerabilities. During testing, several types of vulnerabilities were introduced into the code to verify the tool's effectiveness. The tool successfully identified insecure code and provided warnings. Below are some test results displayed by vs code IDE.

A. Cross Site Scripting Detection

Scenario: Code that inserts user input into the HTML content using **innerHTML**, which is susceptible to XSS attacks.

```
JS xss_test.js > ...
1
2   const Input1 = userInput1;
3   document.getElementById('output').innerHTML = Input1;
4
5   document.getElementById('output').textContent = Input1;
6
7   document.getElementById('output').innerHTML = DOMPurify.encodeURI(Input1);
8
```

Fig. 5. HTML Code

Observation: The extension(tool) highlighted the insecure usage of *innerHTML* and warned about the XSS risk. It recommended using safer alternatives such as ***textContent*** or sanitizing the input.

B. SQL Injection detection

Scenario: An SQL query constructed using un-sanitized user inputs, leading to SQL Injection vulnerabilities.

```
//Basic Injection
//vulnerable
app.post('/login', (req, res) => {
    var username = req.body.username;
    var password = req.body.password;
    var sql = `SELECT * FROM users WHERE username = '${username}' AND password = '${password}'`;

    connection.query(sql, function (error, results, fields) {
        if (error) throw error;
        // Proceed with authentication success logic
        res.send('Logged in successfully');
    });
});
```

Fig.6. SQL Query

Observation: The tool identified the insecure SQL query and warned about the potential SQL Injection risk. It recommended using parameterized queries to prevent this vulnerability.

C. Cross-Site Request Forgery Detection

Scenario: A sample API endpoint that updates user profiles without CSRF protection.

```
// Normally updating user profile without CSRF protection
app.post('/api/update-profile', (req, res) => {
    const profileData = req.body;
    console.log('Profile Data:', profileData);
    res.json({ message: 'Profile updated successfully!' });
});
```

Fig.7. Sample API endpoint

Observation: The tool detected the lack of CSRF protection and warned about the associated risk. It recommended implementing CSRF tokens.

The testing and validation phase confirmed that the VS Code extension made using Static Analysis meets its functional requirements and provides a robust solution for enhancing web application security during development. By detecting and mitigating vulnerabilities in real-time, the tool helps developers adhere to secure coding practices and reduces the risk of security breaches in web applications. Other than the detected vulnerabilities the algorithms can be implemented to detect various vulnerability types as the vulnerability database being updated. Not only detection of vulnerabilities, Static Analysis made the detection more fasrer when compared to Dynamic Analysis in existing tools. The integration to the IDE via extension it made the tool to run real-time with the development environment. Not only that the extension has the

potential to significantly impact how new developers are introduced to security practices, making secure coding more accessible. Future enhancements could include expanding the range of detectable vulnerabilities,

supporting more programming languages, and possibly revisiting the integration of machine learning for advanced pattern recognition as well.

REFERENCES

[1] Ö. Aslan, S. S. Aktuğ, M. Ozkan-Okay, A. A. Yilmaz, and E. Akin, "A Comprehensive Review of Cyber Security Vulnerabilities, Threats, Attacks, and Solutions," *Electronics (Switzerland)*, vol. 12, no. 6. MDPI, Mar. 01, 2023. doi: 10.3390/electronics12061333.

[2] D. Kaur and P. Kaur, "Empirical Analysis of Web Attacks," in *Physics Procedia*, Elsevier B.V., 2016, pp. 298–306. doi: 10.1016/j.procs.2016.02.057.

[3] V. P. Dwivedi Scholar, C. Rits, B. Himanshu Yadav Asst, C. dept RITS, and B. Anurag Jain, "Web Application Vulnerabilities: A Survey," 2014.

[4] F. M. Tudela, J. R. B. Higuera, J. B. Higuera, J. A. S. Montalvo, and M. I. Argyros, "On combining static, dynamic and interactive analysis security testing tools to improve owasp top ten security vulnerability detection in web applications," *Applied Sciences (Switzerland)*, vol. 10, no. 24, pp. 1–26, Dec. 2020, doi: 10.3390/app10249119.

[5] L. Dencheva and M. Monaghan, "Comparative analysis of Static application security testing (SAST) and Dynamic application security testing (DAST) by using open-source web application penetration testing tools MSc Internship MSc in Cyber Security." [Online]. Available: https://www.enisa.europa.eu/publications/enisa- threat-landscape-2021

[6] S. Motogna, D. Cristea, D. Şotropa, and A. J. Molnar, "Formal concept analysis model for static code analysis," *Carpathian Journal of Mathematics*, vol. 38, no. 1, pp. 159–168, 2022, doi: 10.37193/CJM.2022.01.13.

[7] B. Loring, D. Mitchell, and J. Kinder, "Sound regular expression semantics for dynamic symbolic execution of JavaScript," in *Proceedings of the ACM SIGPLAN Conference on Programming Language Design and Implementation (PLDI)*, Association for Computing Machinery, Jun. 2019, pp. 425–438. doi: 10.1145/3314221.3314645.

[8] L. Groner and G. Manricks, *JavaScript regular expressions : leverage the power of regular expressions to create an engaging user experience*.

[9] Tony Stubblebine, "Regular Expression Pocket Reference," Sebastopol, 2007.

[10] M. Johns, "Code-injection vulnerabilities in web applications - Exemplified at cross-site scripting," *IT - Information Technology*, vol. 53, no. 5, pp. 256–260, Sep. 2011, doi: 10.1524/itit.2011.0651.

[11] Stuttard, Dafydd, Pinto, and Marcus, "The Web Application Hacker's Handbook."

Integrative Approach to Cardiovascular Health Management

Roland Appuhamy
Faculty of Computing
Sri Lanka Institute of Information Technology
Malabe, Sri Lanka
rsdew.appuhami99@gmail.com

Tharindu Lakshan
Faculty of Computing
Sri Lanka Institute of Information Technology
Malabe, Sri Lanka
lakshanjat1@gmail.com

Oshadha Amaraweera
Faculty of Computing
Sri Lanka Institute of Information Technology
Malabe, Sri Lanka
oshadageeth@gmail.com

Parami Jayasekara
Faculty of Computing
Sri Lanka Institute of Information Technology
Malabe, Sri Lanka
paraminawodya500@gmail.com

Kapila Dissanayaka
Faculty of Computing
Sri Lanka Institute of Information Technology
Malabe, Sri Lanka kapila.d@sliit.lk

***Abstract*— Heart diseases remain a significant global health challenge, emphasizing the importance of early intervention and prevention. In addressing this issue, a mobile application was developed with four main components focusing on predicting the risk of becoming a heart patient, predicting the risk category of a heart patient, food item identification and dietary recommendation, and suggesting best rated nearby pharmacies to fulfill prescriptions. All these components include trained Artificial Intelligence models to assure a reliable performance within each component. In predicting the risk of becoming a heart patient, a trained logistic regression algorithm was used. As for predicting the risk category of a heart patient, a trained K-Nearest Neighbors algorithm was used. For the components of food item classification and suggesting best rated nearby pharmacies, YOLOv5 and TextBlob were utilized respectively. Results show that the accuracies of the trained models in the components are significant enough for mass scale utilization of all the four components of our mobile application in improving life standards of heart patients.**

Keywords—Cardiac health prediction, predictive modeling, logistic regression, machine learning algorithm, YOLOv5

I. INTRODUCTION

Heart diseases remain a leading cause of mortality with cardiac arrest posing a significant threat to public health. Despite remarkable advancements in medical science, accurately predicting the occurrence of a heart attack and its subsequent trajectory remains a critical challenge.

Early prediction of heart attacks and forecasting the likelihood of future occurrences are imperative for effective intervention and management. Furthermore, individuals who have suffered heart attacks require vigilant monitoring and access to appropriate medical resources to mitigate further risks and ensure optimal recovery. Therefore, it is necessary to develop predictive models capable of identifying individuals at high risk of heart attacks and forecasting the likelihood of subsequent occurrences.

The complexity of heart disease necessitates sophisticated approaches for accurate prediction and prevention. However, existing predictive models often lack the granularity and specificity required for effective risk assessment. Hence, it is better to suggest a comprehensive solution that empowers individuals to make informed dietary choices and access necessary medications promptly. By addressing these concerns positively, we can enhance the efficacy of

preventative measures and improve outcomes either for individuals who are at a high risk of having heart

attacks or for individuals who are in the recovery phase after being a victim of a heart attack.

In addressing the above concerns, we developed an easy- to-use mobile app that can serve as a hub to handle the four functional components: predicting the risk of becoming a heart patient, predicting the risk category of a heart patient, providing food identification and dietary recommendation, and suggesting best rated nearby pharmacies to fulfill prescriptions. This approach guarantees a convenient and accurate way for individuals to assess cardiovascular health frequently while providing insights into mitigatory actions for not to become a victim of sudden heart failure.

II. LITERATUE REVIEW

Frequent cardiovascular health assessment is highly recommended for heart patients to take necessary actions before it becomes life threatening to them. In doing so, Artificial Intelligence (AI) or Machine Learning (ML) based predictive modeling plays a vital role since prediction accuracy can be tuned to trigger higher patient outcomes. In a study published in 2018 shows that AI is revolutionizing healthcare by automating diagnosis, improving data quality, reducing patient care variability, and lowering costs. They focus on AI predictive models for chronic disease diagnosis, enhancing early detection and patient outcomes. The majority of studies used regression-based ML models for diagnosing Celiac Disease, with 45% using Support Vector Machine (SVM), 23% using K-Nearest Neighbors (KNN) and Naïve Bayes (NB) models, 18% using Linear Regression, and 14% using Random Forest (RF) models. RF models were particularly effective in predicting liver fibrosis stages and identifying the liver fibrosis index degree. Bayesian network models were also used to analyze Chronic Obstructive Pulmonary Disease (COPD) patient data. SVM models showed better accuracy in predicting exacerbations and COPD detection [1]. AI can also revolutionize cardiovascular imaging since it can improve diagnostic accuracy, and physician productivity. However, challenges like interpretability and potential harmful recommendations remain [2]. In the study that comes under [3], they review the use of machine learning in predicting and early detection of cardiovascular issues in cancer survivors considering cardiac dysfunction and ECG use. It also highlights the prospects of

ML in cardio-oncology. In another study, they explore the use of traditional machine learning algorithms and advanced technologies for predicting cardiovascular disease risk estimation. It compares various classifiers with J48 achieving about 70% accuracy and Keras achieving about 80% accuracy [4]. In their study [5], they highlight that heart diseases are a major global mortality cause, with ECGs being the most cost- effective diagnostic tool. However, challenges like limited experts and complex interpretations hinder accurate diagnosis. In this scenario, machine learning can be incorporated to unravel those complex interpretations along with the full explanations of the machine leaning models so that the practitioners can trust the results.

According to the study in [6], AI models are transforming medicine, enhancing efficiency and allowing medical professionals to focus on complex tasks. AI models can be used for disease prediction, diagnosis, and decision-making, improving data quality, reducing costs, and forecasting chronic diseases. Model accuracy in various diseases is a critical factor in predicting outcomes. They highlight that the accuracy of their models on predicting diabetes can vary in the range of 73.1-91.6%. The models that predict cardiac diseases show an accuracy within the range of 84-91%. NB, RF, KNN, and SVM models have the highest accuracy in liver disease prediction. LR models have been effective in identifying depression reasons, while KNN models are better for identifying disease patterns. The study in [7] proposes a data-driven approach using machine learning and electronic medical records to reduce preventable hospital readmissions for heart failure, stroke, and pneumonia. The Naïve Bayes algorithm outperformed existing models, achieving an accuracy of 83.19% and an Area Under Curve (AUC) of 0.78 with the possibility to enforce further refinement and validation [7]. Correct and effective assessment of diabetes is crucial to having a healthier heart. To address this fact, a study has been conducted to compare six machine learning algorithms for early diabetes prediction on a medical dataset, aiming to identify the most effective model for improving diagnostic accuracy and aiding doctors and healthcare practitioners in making timely decisions [8]. In 2019, a study has been conducted on a cloud-based healthcare application for predicting heart disease using the University of California Irvine ML Repository dataset and healthcare sensors. Various classification algorithms, including J48, were tested, with the J48 classifier showing the highest accuracy, precision, recall, F-score, and kappa value [9]. Also, in another study, it highlights the potential of machine learning in cancer diagnosis and prediction, but challenges remain in managing large datasets and unstructured information. Efficient machine learning

approaches could revolutionize healthcare by providing precise insights into symptoms, diagnoses, procedures, and treatments [10].

It is clearly noted that all the research mentioned above show the way that AI can be useful in predicting or assessing various diseases including cardiovascular diseases. However, an integrative approach in predicting, suggesting mitigatory actions, and aiding in dietary recommendations and medications are normally not addressed in single research itself. Hence, in this research, we suggest an AI base integrative approach to manage cardiovascular health by providing the features mentioned above with high accuracy.

III. METHODLOGY

In our research, we designed a mobile app for the convenience of heart patients and doctors focusing on risk assessment and control of cardiac failures. Here, the risk of a person becoming a heart patient can be found earlier. And a person who is at risk of heart disease should be advised to follow a dietary plan. Another unique feature here is the ability to automatically analyze the ECG reports and predict the risky situations that a person may face in the future with the help of previously obtained reports. Also, a patient can find the nearest pharmacy to get the prescription through this mobile application as well. The overall system diagram is shown in Fig. 1 and a brief description of the four main functional components is mentioned subsequently.

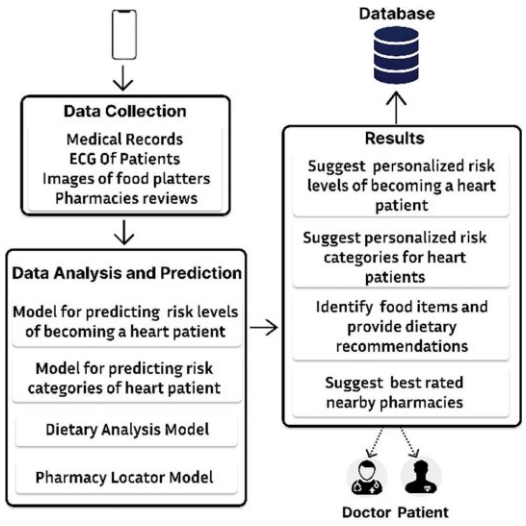


Fig. 1. Overall system diagram for cardiovascular health management.

A. Predicting the Risk Level of Becoming a Heart Patient

The methodology for developing a system for predicting the probability of becoming a heart patient involves several key steps. Firstly, a comprehensive survey was conducted among doctors from selected Sri Lankan government hospitals including Maharagama Apeksha Hospital, Ragama Hospital, Monaragala Sirigala Hospital to understand current diagnostic challenges and delays. Subsequently, data collection encompassed various demographic and health- related factors, including age, sex, cholesterol levels, blood pressure, heart rate, diabetes status, family history, smoking habits, obesity, alcohol consumption, exercise frequency, dietary habits, previous heart problems, medication use, stress levels, sedentary behavior, income, BMI, triglyceride levels, physical activity, and sleep duration. Data from multiple sources, including medical records and surveys, were collected and organized for analysis. This dataset is then preprocessed, including handling missing values, normalization, and feature engineering. Next, a logistic regression model was trained using the collected data, with feature selection and tuning performed to optimize predictive accuracy. The model's performance was evaluated through cross-validation and external validation on independent datasets to ensure robustness and generalizability. The flowchart for this component is shown in Fig. 2.

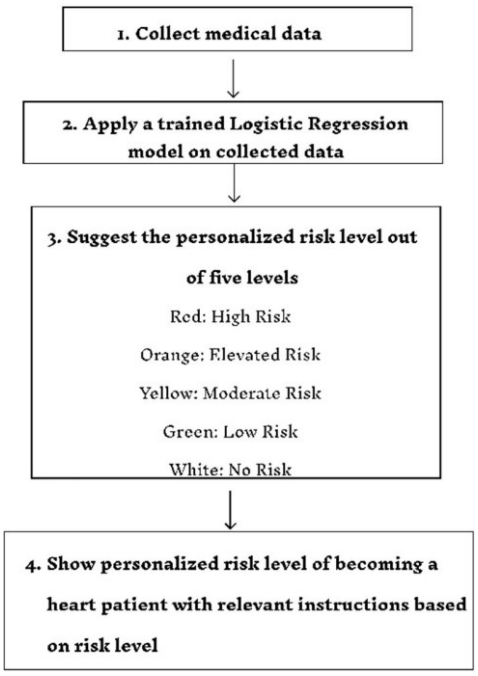


Fig. 2. Flowchart for risk levels of becoming a heart patient.

Subsequently, the model outputs were stratified into five risk levels (High Risk, Elevated Risk, Moderate Risk, Low Risk, and No Risk) according to WHO National Guidelines, enabling clear communication of individual risk status. Instructions for risk management and prevention were developed based on these risk levels, aligning with established guidelines to ensure clinical relevance and effectiveness. Finally, a user-friendly interface was designed to present the results in an easily understandable format, enhancing accessibility for both patients and healthcare providers. Iterative feedback and testing were conducted to refine the system's usability and efficacy, culminating in a comprehensive tool for predicting the possibility of becoming a heart patient and then providing actionable instructions for risk management.

B. Automated Analysis of ECG and Predictive Modeling for Future Heart Attack Risk

This component was developed using machine learning with expert knowledge to analyze ECG records. The required data was obtained from Maharagama Apeksha Hospital. Here, ECG is classified under four categories. They are Normal Person Heartbeat, Abnormal Heartbeat, History of Myocardial Infraction, and Myocardial Infraction. The first step is extracting 12 different leads of the ECG signal from the original ECG image. Each lead is extracted by defining a rectangular region of interest on the image using the slicing notation. Each lead image is then converted to a grayscale image, and the image is then smoothed to remove high-frequency noise. Those twelve leads are combined to create the 13th lead (see Fig. 3) and extract the signal from the Lead- 13 and save it as an image file, which is then ready to be used for further analysis, such as feature extraction and filtering. Then, a comma-separated values (CSV) file is created using the x and y coordinates of the waves in the thirteenth lead. A model was trained using K-Nearest Neighbors machine learning technique using the same CSV file. Here, x and y coordinates are used as input, and the ECG categorical group

as mentioned above is identified as output. The four categories mentioned at the beginning are classified under three risk levels, and a graph shows variation in them. The flowchart for this component is shown under Fig. 4.

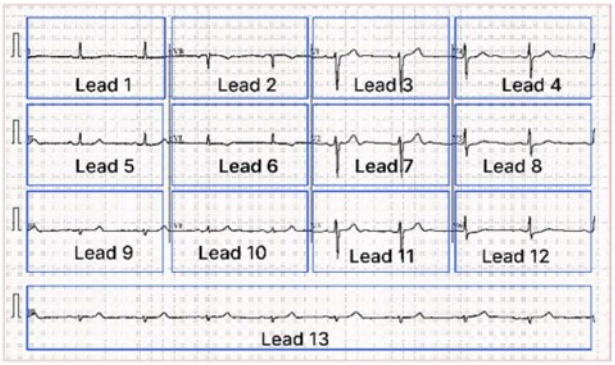


Fig. 3. Sample ECG with all thirteen leads.

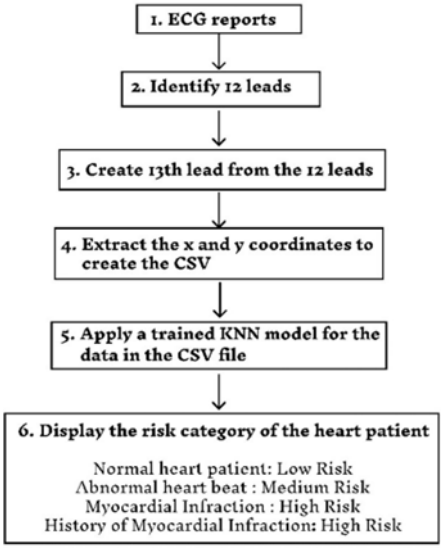


Fig. 4. Flowchart for automated analysis of ECG and predictive modeling.

C. Dietary Guidance System

The process of developing dietary guidance involves the following main steps. The first step is gathering a large amount of food images from various sources such as the internet and pictures manually captured from the camera, which forms the basis of the training dataset for the model. A test model is built and refined using approximately sixty images from reliable online repositories. After the model is validated, it is modified to account for the unique characteristics of Sri Lankan food, which requires a careful curation of one hundred images from the categories of meat, fish, rice, carrots, beetroot, eggs, and potatoes.

Every image in the dataset has been carefully annotated using expert annotation tools such as

"MakeSense" to guarantee accurate labeling and efficient model comprehension. YOLOv5 architecture is used for model training, and Google Colab's processing power is utilized to maximize effectiveness and performance. The flowchart for this component is shown in Fig. 5.

After being integrated into the backend system, the trained model works in unison with the main features of the application, correctly recognizing the food items that are shown and offering relevant dietary recommendations based on the profiles of specific users. This essential feature enables users to make well-informed dietary choices that are in line with their individual nutritional needs and health goals.

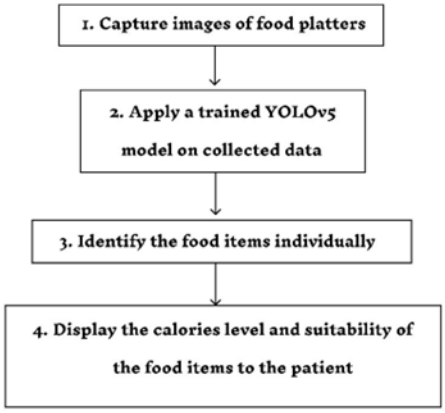


Fig. 5. Flowchart for dietary guidance.

D. Identify the Nearest Pharmacies with the Best Customer Recommendation

The overview of this model is described in Fig. 6, which facilitates different functions that are provided through the developed app. The input data for the model are user location, distance to nearby pharmacies obtained through Google Maps API, availability of required medicines and the ratings. Based on the distance, availability of medicines and the ratings, the pharmacies are listed upon the request of the user for an optimized result. Note that all these functions are performed within the main mobile application.

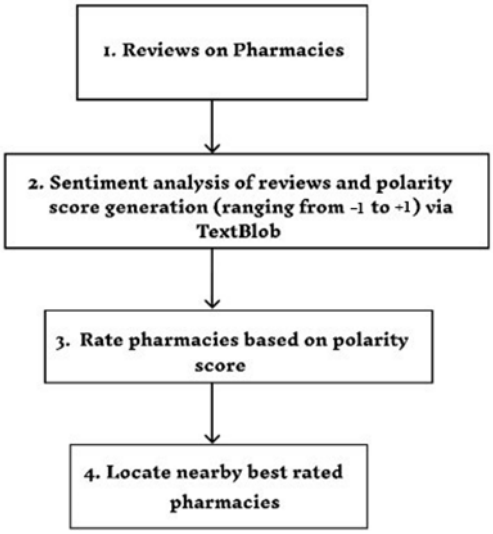


Fig. 6. Flowchart for generating the best rated nearby pharmacies.

The app creates a special token ID that is saved in Firebase and allows the device's token ID to be used to retrieve local pharmacies using Google Places API. Using Google APIs, GeoFire finds pharmacies within a given radius and sorts them automatically. While sentiment analysis uses rules and requires a pre-defined list of classified words, commonly supplied from databases like NLTK, TextBlob uses a machine learning approach to predict scores from user reviews. Sentiment analysis produces more accurate results by identifying sentiments in the input text based on word frequency and semantic correlations. TextBlob, a pre-trained Python library, is used to analyze the sentiment of patients' reviews. It determines whether a review is positive or negative by calculating its polarity score. This score ranges from □1.0 (indicating a highly negative sentiment) to +1.0 (indicating a highly positive sentiment). Based on this polarity score, reviews are classified as positive or negative. Pharmacies can then be ranked or positioned according to the average polarity score of their associated reviews. Based on this score, a corresponding sentiment rating will be assigned to each review. To display the higher rated Pharmacies, Multinomial Naïve Bayes algorithm is applied after vectorization of the reviews.

IV. RESULTS AND DISCUSSION

The introduction of the proposed system shows several significant challenges faced by doctors and patients when managing health. The probability of having a heart attack was predicted through logistic regression model with training accuracy of 94% and testing accuracy of 93%. Stratification of predicted probabilities into five risk levels adheres to WHO

National Guidelines, facilitating clear risk communication. Utilizing WHO National Guidelines, the model stratified predicted probabilities into five distinct risk levels: Red (High Risk), Orange (Elevated Risk), Yellow (Moderate Risk), Green (Low Risk), and White (No Risk). These risk levels facilitate clear communication of the individual's heart health status. Additionally, the system provides relevant instructions based on risk levels by considering WHO National Guidelines, ensuring appropriate actions and interventions are recommended for each risk category. The user-friendly interface enhances accessibility, aiding both patients and healthcare providers in understanding and navigating the results effectively. The integration of WHO National Guidelines into the logistic regression model represents a significant step forward in personalized cardiovascular risk assessment. By combining accuracy, standardization, and practicality, this approach holds promise for improving the early identification and management of individuals at risk of heart attacks, ultimately contributing to better heart health outcomes on a global scale.

Users of the mobile application can upload images of their meals to accurately identify the food items with an accuracy of around 85%. This has been tested only for a particular set of common foods used by Sri Lankans such as rice, potato, beetroot, fish, eggs, etc. By utilizing advanced algorithms, the system offers comprehensive details about every item, such as its calories content and suitability for specific patient conditions. The tailored approach makes it easier for users to make educated dietary decisions that meet their nutritional needs and overall health goals.

The KNN model, which was used in predicting the risk level of having a future heart attack, has a training accuracy of 96%. However, the testing accuracy drops to 88%. This could be because the ECG reports are unique from patient to patient. Note that the model effectively analyzes the ECG images to derive the ECG category and the risk level from that category. The normal person category was grouped as low risk (represented by 0), the abnormal heartbeat category

as medium risk (represented by 1), and the myocardial infraction and history of myocardial categories were selected as high risk (represented by 2). Accordingly, a graph generated by the mobile application is shown below (see Fig. 7) with Risk Level in y-axis (vertical) and Dates in x- axis.

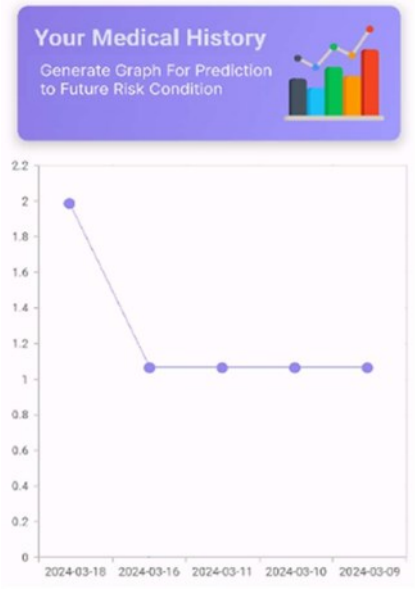


Fig. 7. Sample risk category levels of having a heart attack.

Identifying the nearest pharmacies with the best customer recommendations was implemented successfully. A sample suggesting the best rated pharmacies—within a given radius via google map—on user request is displayed in Fig. 8.

Fig. 8. Sample list of best rated nearby pharmacies.

V. CONCLUSION

The integrative approach presented in this paper offers a significant step forward in the early prediction and management of cardiovascular health. By integrating the four functional components: predicting the risk of becoming a heart patient, categorizing the risk level for heart patients,

providing personalized dietary recommendations, and locating top-rated nearby pharmacies, the mobile application developed enhances both preventive care and post-heart attack recovery. The user-friendly interface allows individuals to assess their cardiovascular health frequently and take proactive steps towards reducing their risk of having a sudden heart failure. Developing a commercial-grade version of this application holds great potential for serving a large community to provide effective cardiovascular health management and to enhance the quality of life of those individuals who are either at risk of having heart attacks or who are recovering from such situations.

REFERENCES

[1] W. J. Kipp et al., "Artificial Intelligence in Cardiology," Journal of the American College of Cardiology, vol. 71, no. 23, pp. 2668-2679, 2018.

[2] L. P. Badano, D. M. Keller, D. Muraru, C. Torlasco, and G. Parati, "Artificial intelligence and cardiovascular imaging: A win-win combination," (in eng), Anatol J Cardiol, vol. 24, no. 4, pp. 214-223, Oct 2020.

[3] M. Daniel Sierra-Lara et al., "Artificial intelligence opportunities in cardio-oncology: Overview with spotlight on electrocardiography," American Heart Journal Plus: Cardiology Research and Practice, vol. 15, p. 100129, 2022.

[4] M. Ashraf et al., "Prediction of Cardiovascular Disease Through Cutting-Edge Deep Learning Technologies: An Empirical Study Based on TENSORFLOW, PYTORCH and KERAS," Singapore, 2021, pp. 239-255: Springer Singapore.

[5] Y. M. Ayano, F. Schwenker, B. Dufera, and T. Debelee, "Interpretable Machine Learning Techniques in ECG-Based Heart Disease Classification: A Systematic Review," Diagnostics, vol. 13, p. 111, 2022.

[6] G. Battineni, G. G. Sagaro, N. Chinatalapudi, and F. Amenta, "Applications of Machine Learning Predictive Models in the Chronic Disease Diagnosis," (in eng), J Pers Med, vol. 10, no. 2, Mar 31 2020.

[7] K. Shameer et al., "Predictive modeling of hospital readmission rates using electronic medical record-wide machine learning: A case-study using mount sinai heart failure cohort," Pacific Symposium on Biocomputing, vol. 0, no. 212679, pp. 276-287, 2017.

[8] M. A. Sarwar, N. Kamal, W. Hamid, and M. Shah, "Prediction of Diabetes Using Machine Learning Algorithms in Healthcare," pp. 1-6, 2018.

[9] M. Ganesan and N. Sivakumar, "IoT based heart disease prediction and diagnosis model for healthcare using machine learning models," in 2019 IEEE International Conference on System, Computation, Automation and Networking (ICSCAN), 2019, pp. 1-5.

[10] N. B and I. Velchamy, "Predictive analytics in health care using machine learning tools and techniques," pp. 492-499, 2017.

A method of directly defining the inverse mapping for solutions of cauchy reaction-diffusion problems

B.S.K. De Silva[1*]
Postgraduate Institute of Science
University of Peradeniya
Peradeniya, Sri Lanka
sahankalana95@gmail.com

M.T.M. Dewasurendra[2]
Department of Mathematics
Faculty of Science
University of Peradeniya
Peradeniya, Sri Lanka
mangalagama.dewasurendra@sci.pdn.ac.lk

***Abstract*— In this paper, a semi-analytical approach called the Method of Directly Defining inverse Mapping (MDDiM) is applied to obtain an approximation series solution for the time-dependent reaction-diffusion problems. This novel approach is being applied to reaction-diffusion problems for the first time. Reaction-diffusion equations are used to describe, in modeling, nonlinear systems across various fields, including physics, chemistry, ecology, biology, and engineering. Traditionally, perturbation and asymptotic methods were widely used to obtain analytical approximations for nonlinear problems. However, these methods often failed under strong nonlinearity conditions. To overcome this obstacle, Shijun Liao introduced the Homotopy Analysis Method (HAM) in 1992, which is effective for highly nonlinear problems. However, determining the unknown functions requires computing the inverse of the linear operator. In scientific computing, calculating the inverse operator for a differential equation is a time-consuming process. To overcome this obstacle, in 2016 Liao introduced a new approach, namely the Method of Directly Defining inverse Mapping (MDDiM) with the freedom to choose the inverse linear mapping directly. The proposed method achieved a remarkable level of accuracy for solutions even with few iterations. Therefore, it can be concluded that MDDiM is not only easy to use but also accurate.**

Keywords— ***Cauchy reaction-diffusion problems, Method of directly defining inverse mapping, Homotopy analysis method***

I. INTRODUCTION

One-dimensional time-dependent reaction-diffusion equation is expressed as follows:

$$\frac{\partial u}{\partial t}(x,t) = D\frac{\partial^2 u}{\partial x^2}(x,t) + r(x,t)u(x,t), \quad (x,t) \in \Omega \subset \Re^2, \tag{1}$$

where u is the concentration, r is the reaction parameter and $D > 0$ is the diffusion coefficient. The initial and boundary conditions are

$$u(x,0) = f(x), \quad x \in \Re, \tag{2}$$

$$u(0,t) = g_0(t), \quad \frac{\partial u}{\partial x}(0,t) = g_1(t), \quad t \in \Re. \tag{3}$$

Reaction-diffusion equations are used to describe nonlinear systems in fields such as physics, chemistry, ecology, biology, and engineering [1-4]. Lesnic [5] provided approximate series solutions for reaction-diffusion problems using the analytic Adomian Decomposition Method (ADM). Dehghan and Shakeri [6] also employed the Variational Iteration Method (VIM) to obtain similar approximate solutions for these problems. Bataineh, Noorani, and Hashim [7] used the Homotopy Analysis Method (HAM) to solve Cauchy reaction-diffusion problems. In 2016, the Method of Directly Defining inverse Mapping (MDDiM) was introduced by Liao [8] to solve a nonlinear ordinary differential equation. The Method of Directly Defining inverse Mapping builds upon the Homotopy Analysis Method [9-12], an analytical approximation technique used for solving highly nonlinear differential equations. This method allows the freedom to choose the inverse linear mapping directly, avoiding the need to calculate the inverse of an operator. Recently, the MDDiM was extended by Vajravelu and his research group [13-17] to systems of nonlinear ordinary differential equations. Also, this novel technique was further developed to solve single and coupled nonlinear PDEs [18-19]. In this study, we further applied this novel method to address Cauchy reaction-diffusion problems. All computational results were achieved using Maple 16.

II. MATERIAL AND METHODS

First of all, we describe the basic ideas of the MDDiM.

Consider the following nonlinear partial differential equation,

$$\mathcal{N}[u(\bar{x})] = 0\,, \quad \bar{x} \in \Omega\,, \tag{4}$$

subject to μ linear boundary conditions

$$\mathfrak{B}_i[u(\bar{x})] = \beta_i\,, \quad \text{at } \bar{x} = \bar{\alpha}_i\,, \qquad i = 1,2,3,\dots,\mu. \tag{5}$$

where $\mathcal{N}$ is a nonlinear operator, $u(\bar{x})$ is an unknown function, $\bar{x}$ denotes the independent variable, Ω is the domain of $\bar{x}$, $\mathfrak{B}_i$ denotes a linear operator, $1 \le \mu \le n$ are positive integers, $\bar{\alpha}_i \in \Omega$ and $\beta_i (1 \le i \le \mu)$ are constants, n is the order of (4), respectively.

Let $S_\infty = \{\varphi_1(\vec{x}), \varphi_2(\vec{x}), \dots\}$ denote a complete set of an infinite number of base functions that are linearly independent. All functions that are expressed by S_∞ form a set of functions, denoted by $V = \sum_{k=0}^{+\infty} a_k \varphi_k(\vec{x})$. The approximate solution comes from this space. Next define the space for the initial guess $V^* = \sum_{k=0}^{\mu} a_k \varphi_k(\vec{x})$, taking the linear combination of first μ functions of the set S_∞, and $\hat{V} = \sum_{k=\mu+1}^{+\infty} b_k \varphi_k(\vec{x})$ so that $V = V^* + \hat{V}$. Assume that $u(\vec{x}) \in V$ and the μ unknown coefficients can be uniquely determined by the μ linear boundary conditions (5). Similarly, letting $S_R = \{\psi_1(\vec{x}), \psi_2(\vec{x}), \dots\}$ and then define $U = \sum_{k=0}^{+\infty} c_k \psi_k(\vec{x})$ so that $\mathcal{N}[u(\vec{x})] \in U$. Finally directly defining the inverse mapping $\Phi: U \to V$. We obtain the following higher order deformation equations for the MDDiM [8,19]:

$$u_k(\vec{x}) = \chi_k u_{k-1}(\vec{x}) + c_0 \Phi[\delta_{k-1}(\vec{x})] + \sum_{n=1}^{\mu} a_{k,n} \varphi_n(\vec{x}). \quad (6)$$

For the error analysis, we first consider the n-term solution (or the approximate series solution) $u(\vec{x}) = u_0(\vec{x}) + \sum_{k=1}^{n-1} u_k(\vec{x})$, using (6) and choosing an appropriate initial guesses $u_0(\vec{x})$. Then, by adding the first n terms of the series solution, we define square residual error function $E(c_0)$ to find the c_0 value which gives the optimal errors, as

$$E(c_0) = \int_\Omega (\mathcal{N}[u(\vec{x})])^2 d\vec{x}. \quad (7)$$

However, in practice, the evaluation of $E(c_0)$ is much more time-consuming. Therefore, instead of using the exact residual error, we use the average (or discrete) residual error, defined as:

$$E(c_0) = \frac{1}{(M+1)(N+1)} \sum_{i=0}^{M} \sum_{j=0}^{N} (\mathcal{N}[\hat{u}(\frac{i}{M+1}, \frac{j}{N+1})])^2. \quad (8)$$

The suggested method provides a very precise and dependable solution for even relatively few terms. Also, MDDiM always uses only less CPU time. Further, it is important to note that finding an inverse linear operator that works well (gives a low error or leads to easily generated solution terms, or both) for a particular type of problem would be worth investigating.

III. Results and Discussion

In this section, we illustrate the accuracy of the MDDiM for various cases of $r(x, t)$. The selected examples are those for which analytical solutions are already known.

A. Example 1. Case r = constant

Taking $D = 1$ and $r = -1$, problem (1)-(3) becomes:

$$\frac{\partial u}{\partial t}(x,t) = \frac{\partial^2 u}{\partial x^2}(x,t) - u(x,t), \quad (x,t) \in \Omega \subset \Re^2, \quad (9)$$

subject to the initial and boundary conditions

$$u(x,0) = e^{-x} + x = f(x), \quad x \in \Re, \quad (10)$$

$$u(0,t) = 1 = g_0(t), \quad (11)$$

$$\frac{\partial u}{\partial x}(0,t) = e^{-t} - 1 = g_1(t), \quad t \in \Re. \quad (12)$$

Then, it was straight forward to choose initial guess $u_0(x,t) = e^{-x} + x$. In the frame work of the MDDiM, we directly defined the inverse mapping, $\Phi: U \to V$

$$\Phi[t^k] = \frac{t^{k+1}}{Ak+1} \quad (13)$$

where A is constant to be selected. The special solutions is

$$u_k(x,t) = \chi_k u_{k-1}(x,t) + c_0 \Phi[\mathcal{D}_{k-1}] + a_{k,0}, \quad (14)$$

where $a_{k,0}$ determined by $u_k(x,0) = 0$.

Using only five terms, let $u(x,t) = u_0(x,t) + u_1(x,t) + u_2(x,t) + u_3(x,t) + u_4(x,t)$. The sum of squared residual error functions is given by

$$E(c_0, A) = \frac{1}{250000} \sum_{i=0}^{499} \sum_{j=0}^{499} (\mathcal{N}[u(\frac{i}{500}, \frac{j}{500})])^2. \quad (15)$$

Note that $E(c_0, A)$ is a function of c_0 and A.

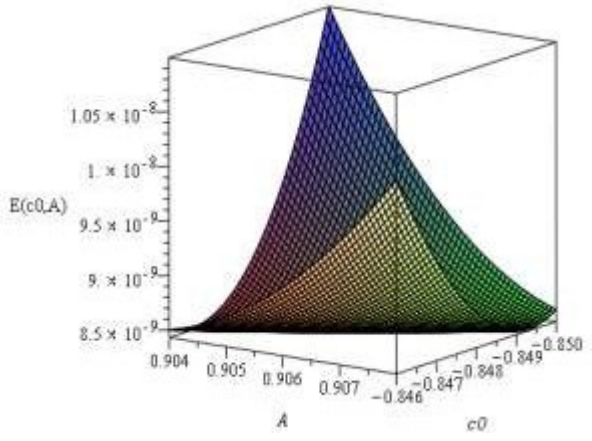


Fig. 1. Plot of $E(c_0, A)$, the squared residual error as a function of c_0 and A for example 1. The error function has minimum $E(c_0, A) = 8.4294 \times 10^{-9}$ where $c_0 = -0.8480$ and $A = 0.9062$.

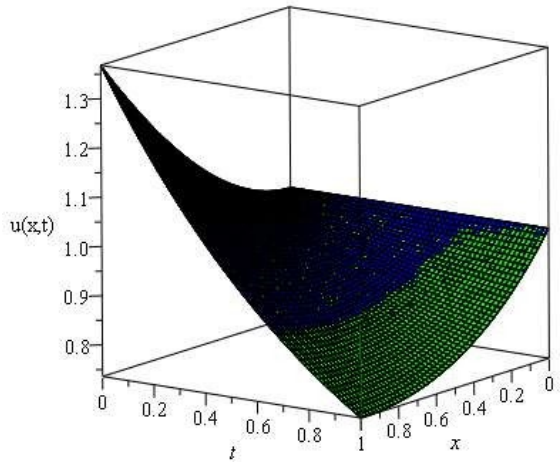


Fig. 2 Graph of the approximations solution and exact solution for example 1.

B. Example 2. Case $r = r(t)$

Taking $D = 1$ and $r(t) = 2t$, problem (1)-(3) becomes:

$$\frac{\partial u}{\partial t}(x,t) = \frac{\partial^2 u}{\partial x^2}(x,t) + 2tu(x,t), \quad (x,t) \in \Omega \subset \Re^2, \tag{16}$$

subject to the initial and boundary conditions

$$u(x,0) = e^x = f(x), \quad x \in \Re, \tag{17}$$

$$u(0,t) = e^{t+t^2} = g_0(t), \tag{18}$$

$$\frac{\partial u}{\partial x}(0,t) = e^{t+t^2} = g_1(t), \quad t \in \Re. \tag{19}$$

Then, it was straight forward to choose initial guess $u_0(x,t) = e^x$. In the frame work of the MDDiM, the same inverse mapping (13) was used.

The special solutions is

$$u_k(x,t) = \chi_k u_{k-1}(x,t) + c_0 \Phi[\mathcal{D}_{k-1}] + a_{k,0}, \tag{20}$$

where $a_{k,0}$ determined by $u_k(x,0) = 0$.

Using only five terms, let $u(x,t) = u_0(x,t) + u_1(x,t) + u_2(x,t) + u_3(x,t) + u_4(x,t)$. The sum of squared residual error functions is given by

$$E(c_0, A) = \frac{1}{250000} \sum_{i=0}^{499} \sum_{j=0}^{499} \left(\mathcal{N}\left[u\left(\frac{i}{500}, \frac{j}{500}\right)\right]\right)^2. \tag{21}$$

Note that $E(c_0, A)$ is a function of c_0 and A.

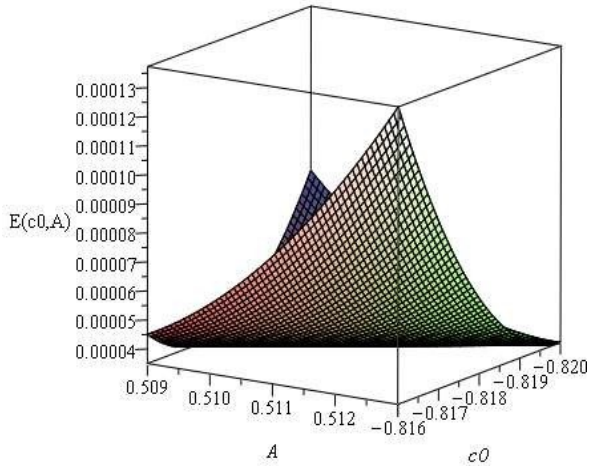


Fig. 3. Plot of $E(c_0, A)$, the squared residual error as a function of c_0 and A for example 2. The error function has minimum $E(c_0, A) = 3.5003 \times 10^{-5}$ where $c_0 = -0.8189$ and $A = 0.5113$.

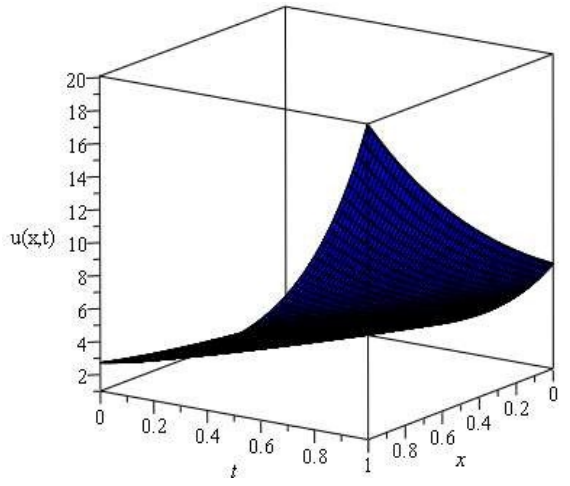


Fig. 4. Graph of the approximations solution and exact solution for example 2.

C. Example 3. Case $r = r(x)$

Taking $D = 1$ and $r(x,t) = -1 - 4x^2$, problem (1)-(3) becomes:

$$\frac{\partial u}{\partial t}(x,t) = \frac{\partial^2 u}{\partial x^2}(x,t) - (1 + 4x^2)u(x,t), \quad (x,t) \in \Omega \subset \Re^2, \tag{22}$$

subject to the initial and boundary conditions

$$u(x,0) = e^{x^2} = f(x), \quad x \in \Re, \tag{23}$$

$$u(0,t) = e^t = g_0(t), \tag{24}$$

$$\frac{\partial u}{\partial x}(0,t) = 0 = g_1(t), \quad t \in \Re. \tag{25}$$

Then, it was straight forward to choose initial guess $u_0(x,t) = e^{x^2}$. In the frame work of the MDDiM, the same inverse mapping (13) was used.

The special solutions is

$$u_k(x,t) = \chi_k u_{k-1}(x,t) + c_0 \Phi[\mathcal{D}_{k-1}] + a_{k,0}, \quad (26)$$

where $a_{k,0}$ determined by $u_k(x,0) = 0$.

Using only five terms, let $u(x,t) = u_0(x,t) + u_1(x,t) + u_2(x,t) + u_3(x,t) + u_4(x,t)$. The sum of squared residual error functions is given by

$$E(c_0, A) = \frac{1}{250000} \sum_{i=0}^{499} \sum_{j=0}^{499} \left(\mathcal{N}\left[u\left(\frac{i}{500}, \frac{j}{500}\right)\right]\right)^2. \quad (27)$$

Note that $E(c_0, A)$ is a function of c_0 and A.

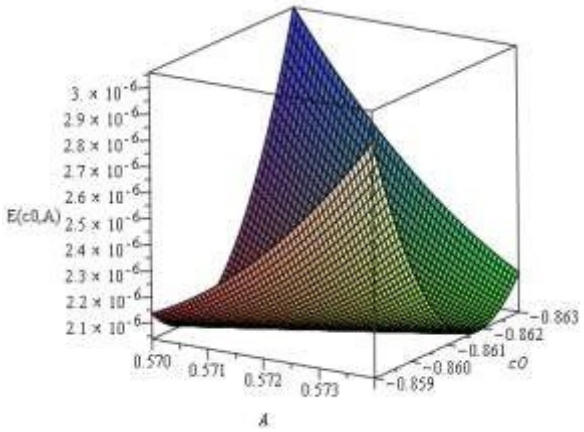


Fig. 5. Plot of $E(c_0, A)$, the squared residual error as a function of c_0 and A for example 3. The error function has minimum $E(c_0, A) = 2.0383 \times 10^{-6}$ where $c_0 = -0.8610$ and $A = 0.5723$.

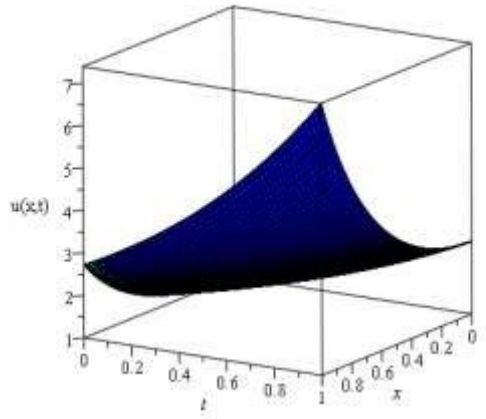


Fig. 6. Graph of the approximations solution and exact solution for example 3.

D. Example 4. Case r = r(x,t)

Taking $D = 1$ and $r(x,t) = -4x^2 + 2t - 2$, problem (1)-(3) becomes:

$$\frac{\partial u}{\partial t}(x,t) = \frac{\partial^2 u}{\partial x^2}(x,t) - (4x^2 - 2 + 2)u(x,t), \quad (x,t) \in \Omega \subset \Re^2, \quad (28)$$

subject to the initial and boundary conditions

$$u(x,0) = e^{x^2} = f(x), \quad x \in \Re, \quad (29)$$

$$u(0,t) = e^{t^2} = g_0(t), \quad (30)$$

$$\frac{\partial u}{\partial x}(0,t) = 0 = g_1(t), \quad t \in \Re. \quad (31)$$

Then, it was straight forward to choose initial guess $u_0(x,t) = e^{x^2}$. In the frame work of the MDDiM, the same inverse mapping (13) was used.

The special solutions is

$$u_k(x,t) = \chi_k u_{k-1}(x,t) + c_0 \Phi[\mathcal{D}_{k-1}] + a_{k,0}, \quad (32)$$

where $a_{k,0}$ determined by $u_k(x,0) = 0$.

Using only five terms, let $u(x,t) = u_0(x,t) + u_1(x,t) + u_2(x,t) + u_3(x,t) + u_4(x,t)$. The sum of squared residual error functions is given by

$$E(c_0, A) = \frac{1}{250000} \sum_{i=0}^{499} \sum_{j=0}^{499} \left(\mathcal{N}\left[u\left(\frac{i}{500}, \frac{j}{500}\right)\right]\right)^2. \quad (33)$$

Note that $E(c_0, A)$ is a function of c_0 and A.

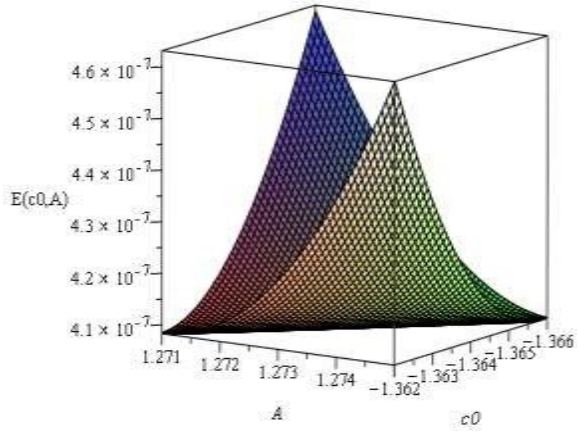


Fig. 7. Plot of $E(c_0, A)$, the squared residual error as a function of c_0 and A for example 4. The error function has minimum $E(c_0, A) = 4.0830 \times 10^{-7}$ where $c_0 = -1.3648$ and $A = 1.2740$.

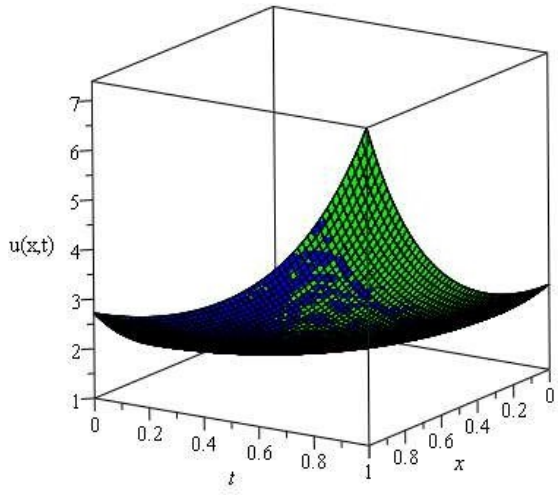


Fig. 8. Graph of the approximations solution and exact solution for example 4.

IV. CONCLUSIONS

In this paper, MDDiM was developed and applied to solve Cauchy reaction-diffusion problems, which are common in various scientific phenomena. Since the direct definition of the inverse operator, the series solutions were obtained using less CPU time, low error, and less complicated terms. The proposed technique produces a highly accurate and reliable solution to the problems in a few iterations. Therefore, it can be concluded that MDDiM is not only easy to use but also accurate.

REFERENCES

[1] N. F. Britton, *Reaction-Diffusion Equations and Their Applications to Biology*. London, U.K.: Academic, 1986, pp. ix–277.

[2] R. S. Cantrell and C. Cosner, *Spatial Ecology via Reaction-Diffusion Equations*. Hoboken, NJ, USA: John Wiley & Sons, 2004.

[3] P. Grindrod, *The Theory and Applications of Reaction-Diffusion Equations: Patterns and Waves*. Oxford, U.K.: Clarendon Press, 1996.

[4] J. Smoller, *Shock Waves and Reaction–Diffusion Equations*, vol. 258. New York, NY, USA: Springer Science & Business Media, 2012.

[5] D. Lesnic, "The decomposition method for Cauchy reaction–diffusion problems," *Appl. Math. Lett.*, vol. 20, no. 4, pp. 412–418, 2007.

[6] M. Dehghan and F. Shakeri, "Application of He's variational iteration method for solving the Cauchy reaction–diffusion problem," *Journal of Computational and Applied Mathematics*, vol. 214, no. 2, pp. 435–446, 2008.

[7] A. S. Bataineh, M. S. M. Noorani, and I. Hashim, "The homotopy analysis method for Cauchy reaction–diffusion problems," *Phys. Lett. A*, vol. 372, no. 5, pp. 613–618, 2008.

[8] S. Liao and Y. Zhao, "On the method of directly defining inverse mapping for nonlinear differential equations," *Numer. Algorithms*, vol. 72, no. 4, pp. 989–1020, 2016.

[9] S. J. Liao, *Homotopy Analysis Method in Non-linear Differential Equations*, 2012.

[10] S. J. Liao, Beyond Perturbation: Introduction to Homotopy Analysis Method, 2003.

[11] K. Vajravelu and R. Van Gorder, *Nonlinear Flow Phenomena and Homotopy Analysis*, vol. 2. Berlin: Springer, 2013.

[12] S. Liao, Ed., *Advances in the Homotopy Analysis Method*. World Scientific, 2013.

[13] M. Dewasurendra, M. Baxter, and K. Vajravelu, "A method of directly defining the inverse mapping for solutions of non-linear coupled systems arising in convection heat transfer in a second grade fluid," *Applied Mathematics and Computation*, vol. 339, pp. 758–767, 2018.

[14] M. Baxter, M. Dewasurendra, and K. Vajravelu, "A method of directly defining the inverse mapping for solutions of coupled systems of nonlinear differential equations," *Numerical Algorithms*, vol. 77, pp. 1199–1211, 2018.

[15] M. Dewasurendra, Y. Zhang, N. Boyette, I. Islam, and K. Vajravelu, "A method of directly defining the inverse mapping for an HIV infection of CD4+ T-cells model," *Applied Mathematics and Nonlinear Sciences*, vol. 6, no. 2, pp. 469–482, 2021.

[16] M. Dewasurendra and K. Vajravelu, "On the method of inverse mapping for solutions of coupled systems of nonlinear differential equations arising in nanofluid flow, heat and mass transfer," *Applied Mathematics and Nonlinear Sciences*, vol. 3, no. 1, pp. 1–14, 2018.

[17] K. Vajravelu, M. Dewasurendra, and Y. Zhang, "A method of directly defining the inverse mapping (MDDiM) for solutions of non-linear coupled systems arising in SIR and SIS epidemic models," *Communications in Numerical Analysis*, vol. 2, pp. 64–77, 2019.

[18] C. W. Sahabandu, D. Karunarathna, P. Sewvandi, Z. A. M. S. Juman, M. Dewasurendra, and K. Vajravelu, "A method of directly defining the inverse mapping for a nonlinear partial differential equation and for systems of nonlinear partial differential equations," *Computational and Applied Mathematics*, vol. 40, no. 6, p. 234, 2021.

[19] C. W. Sahabandu, M. Dewasurendra, Z. A. M. S. Juman, K. Vajravelu, and A. J. Chamkha, "Semi-analytical method for propagation of harmonic waves in nonlinear magneto-thermo-elasticity," *Computers & Mathematics with Applications*, vol. 105, pp. 107–111, 2022.