NSBM GREEN UNIVERSITY TOWN

NSRSIT 2024
Student Research Symposium on Innovative Technology

NSBM

STUDENT RESEARCH SYMPOSIUM ON INNOVATIVE TECHNOLOGY

NSRSIT '24

Organized By,

Faculty of Computing
NSBM Green University

Focus Areas

- **Emerging Technologies** - Explores the latest breakthroughs in fields like AI, Biotechnology, Nanotechnology, and Quantum Computing.
- **Robotics and Automation** - Discusses the impact of digitalization on various industries, including finance, healthcare, and education.
- **Artificial Intelligence and Machine Learning** - Dives deep into Al and ML research, applications, and ethical considerations.
- **Education and EdTech** - Focuses on innovations in education technology, e-learning platforms, and digital pedagogical tools.
- **Blockchain and Cryptocurrency** - Explores the world of blockchain technology, cryptocurrencies, and their potential impact on various industries.
- **Smart Cites and Urban Innovation** - Examines how technology reshapes urban landscapes, transportation, and infrastructure for a more sustainable future.
- **Cyber Security and Data Privacy** - Discusses the latest trends in cybersecurity, data protection, and privacy technologies.
- **Human-Machine Interactions** - Examines research on user interfaces, virtual reality, augmented reality, and human-computer interaction.
- **Cross-disciplinary Collaborations** - Encourages interdisciplinary research and collaboration between different fields of technology and science
- **Extended Reality** - Explores the future of immersive technology with our cutting-edge virtual and augmented reality solutions.

Dean's Message

Dr. Chaminda Wijesinghe
Dean, Faculty of Computing
NSBM Green University

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 inaugural Students' Research Symposium 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 symposium 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, mentors, 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 researchers and thought leaders.

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.

With warmest regards,

Dr. Chaminda Wijesinghe
Dean, Faculty of Computing
NSBM Green University

Organizing Committee

Advisory Board

Prof. EA Weerasinghe, Vice Chancellor, NSBM Green University

Prof. Chaminda Rathnayake, Deputy Vice Chancellor, NSBM Green University

Prof. Baratha Dodankotuwa, Head of Academic development and Quality Assurance

Dr. Chaminda Wijesinghe, Dean, Faculty of Computing

Dr. Rasika Ranaweera, Dean, Faculty of Postgraduate Studies

Prof. Noel Fernando, Senior Lecturer, Faculty of Computing

Symposium Committee

Dr. Mohamed Shafraz, Senior Lecturer, Head/Department of Software Engineering and Information Systems.

Mr. Naji Saravanabavan, Senior Lecturer, Head/Department of Computer Science and Data Science

Mr. Chamara Dissanayake, Senior Lecturer, Dep. of Network and Security

Editorial Committee

Ms. Thisarani Wickramasinghe, Teaching Assistant, Department of Software Engineering and Information Systems.

Ms. Hirushi Dilpriya, Lecturer, Department of Computer Science and Data Science

Ms. Kavishka Rajapaksha, Temporary Lecturer, Department of Computer Science and Data Science

Mr. Diluka Wijesinghe, Teaching Assistant, Department of Software Engineering and Information Systems.

Mr. Anton Jayakody, Teaching Assistant, Department of Computer Science and Data Science

Content

Digital Intervention for Conquering Stress of University 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 counsellors, while questionnaires were distributed among undergraduates. The four 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 the prototype or questionnaire is reviewed. Based on the findings from the Interviews and Questionnaire, both functional and non-functional requirements were gathered. According to the gathered requirements system was designed, taking into account essential UI/UX principles for an optimal user experience.

Keywords—stigma, stressors, stress, academic performance, undergraduates, web application

I. INTRODUCTION

Mental health is defined by the World Health Organization (WHO) as "a state of well-being in which the individual is aware of his or her own abilities, can face the normal stresses of life, can work productively and fruitfully and is capable of making a contribution to their community." Mental health is important because it becomes a part of a human life and its impact on the overall health of the person including physical, mental, and social well-being. It is not just about being sick and showing diseases. Earlier, people didn't talk about mental illness because of their dislike. But now, it must be changed [1]. It costs a lot of money when the person doesn't have proper mental health. According to previous research 12 billion productive working days have been lost due to mental disorders and the cost is nearly US$1trillion [2]. Mental health services are not received 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 from country to country.

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]. About 61% of the surveyed carers held the belief that depression is a sign of personal weakness. Additionally, 60% of the carers believed that a person living with depression is more prone to violence compared to a member of the general community [4].

According to previous research, 12% - 50% of university students have one or more mental disorders [5]. On campuses, mental health issues have grown significantly more important. Every year, over 26% of American adults suffer from diagnosable mental health issues, with many of these illnesses beginning in adolescence. Students in college 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 behaviours, dropping out of college, and even suicidal ideation. Campuses must raise awareness of mental health issues, decrease stigma, offer accessible support services, and put preventive measures in place in order to address these issues. Colleges may support students' academic and personal success by placing a high priority on mental health [6].

Therefore, College students are experiencing an increase in the number of mental health issues in recent years. According to research, around half of university students suffer moderate levels of stress-related mental health conditions such as anxiety and depression [7].

High expectations and demands have been placed on students because of the fast-paced nature of today's world, which is characterized by rapid improvements in technology, education, culture, and society. Within the socio-cultural, economic, and regulatory frameworks of society, these expectations are frequently unclear, unrelated, and unrealistic. As a result, compared to past generations, the present generation of students has much greater levels of stress and anxiety [8].

According to previous research, university students who experience significant levels of psychological distress are troubled with their academics and may not be able to complete their studies. The research also discovered a link between psychological distress and poor academic achievement. Nearly 20% of undergraduate students reported missing academics due to mental health difficulties. Additionally, almost 44% of undergraduate students reported that their academic performance decreased during that undergraduate journey due to mental issues [9].

The Research will be focused on undergraduates who faced stress during their undergraduate time. There has been a concerning rise in the number of undergraduates who show signs of anxiety and stress during the last ten years. Globally, 31% of college students have tested positive for a mental health issue in the previous 12 months. This figure illustrates the strongly increased frequency of mental health issues among undergraduates [10]. The Research further discussed the impact of academic performance as a problem due to stress and the IT intervention to manage stress and increase the academic performance of university students.

A. Research Problem

So, 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. Undergraduate people go through a lot of changes and face many difficulties as they make the transition from being dependent on their family to being more independent throughout their time in college. They may explore their identities, make significant life decisions, and find their purpose and direction [11].

Students in higher education are more likely to have stress and University students have higher rates of mental health issues than the general population, according to research studies like those by the American College Health Association (2018). The collegiate setting offers a variety of pressures that are particular to this developmental stage. Academic constraints that students often deal with include the requirements of coursework, tests, and the desire to maintain excellent marks. They could also struggle with time management, managing many obligations, and adapting to the demanding academic standards and workload of universities [12].

So, it is true that stress among college students might affect their academic progress and even have negative consequences. Stress could affect a student's ability to concentrate, focus, and do well in class. Lower grades, lower rates of progress, or even quitting their academic programs may follow from this.

According to previous research, keeping proper stress management systems physically or digitally can be maintained and all these things can be improved further. Aside from that, involving these systems in academics during the undergraduate period provides an opportunity to facilitate the improvement of academic performance.

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 counsellors as well. It takes time to the treatment. Back in 2018, according to 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 and hopeless, 13% of students considering suicide. Furthermore, that research indicated severity of the mental health and about treatment seeking students [10].

With the technological revolution, counselling services are provided online by developed countries. So, the E-therapist concept has been popular for years. But have a problem what type of E- therapist is wanted by a particular university student.

So, the proposed system is to provide counselling services to the students by providing better services. Basically, what the proposed system does is track stress level of students, considering their degree program and the academic year. While they will be meeting a counsellor, they can do an activity what system is suggested to manage their stress. So that won't take time as well. So, the activities will be sorted according to the type of stress. By reducing stress, students can improve their academic life as well.

B. Research Objectives

The research was conducted based on four main objectives.

1. To identify Academic and Environmental factors for stress of university students which affect their academics.
2. To identify other factors which are caused by stress and problems.
3. To evaluate and study existing IT solutions for stress management.

4. To design the best IT solution for managing stress to improve their academics.

C. *Stress among University Students and Factors Affecting the Stress*

Many students start their college journeys each year with the hope of obtaining degrees that will open the door to successful employment and fulfilled lives. While starting university is a big step for many people and their families, some students may find this period of adjustment to be quite stressful. Along with the difficulties of adjusting to new social dynamics, the transition from the routines of high school to a more independent and challenging academic environment can add to the pressure and emotional strain [13]. Environmental, psychological, biological, and social aspects can all contribute to stress. These elements can have various impacts on people and add to the feelings of stress. The effects of stress can differ based on the severity and length of the stressor, the person's personality attributes, how they think about the stress, and whether they have access to social support. Here we will consider Academic Stress and Environmental Stress which affect the stress of University Students.

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 [14]. 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 [15].

Previous research has mentioned 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 sick, Relationship Issues, and financial issues [16].

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 college students' performance. Students who saw an upcoming exam as a threat were more likely to experience negative emotions and use coping methods centred on managing those emotions rather than directly addressing the stressor. These students who perceived exams 17 as threats and used emotion-focused coping (like trying to calm themselves down) tended to have poorer scores on the exams. This hinted that when students feel overwhelmed by stress, it can negatively impact their performance. Applying Lazarus' theory more broadly, another study found that college students who generally saw academic challenges as threats (like difficult assignments or tests) tended to use less helpful coping methods and experienced more negative emotions.

II. Research Methodology

A. *Research Paradigm*

The research will be followed both positivism and interpretivism. The

positivism approach is being used to collect quantitative data and interpretivism approach is being used to collect qualitative data. To conduct this research, structured interviews, and semi – structured interviews will be conducted. To conduct this research, a semi – structured questionnaire will be conducted, since both qualitative and quantitative data can be gathered. Questionnaires will be shared among private and government university students.

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. Finally, the study seeks to produce a solution to the problem that was brought up in the setting of a Sri Lankan university students. So, the study can be carried out using a deductive reasoning strategy by considering the nature of the research. But in the present research, some of the steps were altered.

C. Research Strategy

Research is conducted by using action research. So, the identified problem of the research is, Academic performance of students is negatively impacted by both academic and environmental stress. By conducting this research as a solution, web application is planning to be implemented by using strategies to manage stress. So, this application can be used at universities. So, the planned intervention can be implemented among undergraduates of the universities. According to the conducted literature review, academic performance is reduced by academic and environmental stress. This will be used as a benchmark for assessing changes following the intervention. Then design and develop the web application using the intended features and approaches.

III. Results and Evaluation

A. Data Collection and Data Analysis

1) Thematic Analysis

The data which is gathered through Interviews is analysed 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.

Themes	**Sub Themes**
Involvement Counsellors for support of students	Existing Communication approaches between undergraduates and counsellors
	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 for 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

Table 1: Themes and Sub Themes

2) Mapping data collection of Questionnaire with results review

The results of the student survey provide a compelling insight into the frequency and severity of academic pressures among the sample group that was examined. According to the findings, a sizeable 58.9% of respondents reported experiencing high levels of stress related to academic expectations. This shows that a substantial proportion of students struggle significantly to meet the demands of their coursework, exams, and other academic requirements. In addition, 27.4% of the students who responded to the survey indicated that they were under moderate stress. This demonstrates that a substantial portion of the student population still experiences considerable amounts of stress because of academic demands. Notably, a majority of respondents,64.4% of the sample surveyed, acknowledged that they struggled with high levels of stress brought on by a variety of environmental circumstances. This emphasizes the significant impact on a student's wellbeing that factors like living arrangements, homesickness, and other environmental stressors can have. Furthermore, 17.8% of the respondents said that environmental stressors had caused them to feel moderately stressed. The primary objective of this research is to improve students' academic performance using efficient stress-reduction techniques. The difference in academic performance levels among the student sample is highlighted by these figures. The Grade Point Average (GPA) is an important measure used to assess this progress. The survey results show that a significant percentage of students have various degrees of academic achievement. It is important to note that 43.8% of students have a GPA of 3.0-3.29 and 32.9% have a GPA of 2.0-2.9. As stress can have a negative effect on GPA, it implies that most students are under some level of stress. According to the survey, 64.4% out of the total respondents have knowledge about the Counselling Service that is present in the university and 44 respondents are not familiar with the counselling service while only 2 undergraduate students are familiar with the counselling service. Therefore, it is evident that even though many respondents had knowledge of counselling services, they are not very familiar with it. This can be a leading cause for the major undergraduates to be under stressed because of the lack of familiarity with the counselling service. According to 98.6% of the undergraduates who completed the survey, undergraduates do not have any knowledge on any existing web – based applications which are designed to help stress management., 78.9 of the undergraduates are neutral on the satisfaction of having a stress management web-based application. This can be due to the undergraduates not having any experience in using a web-based application which helps stress management. Therefore, they cannot give a true opinion on the

system. The majority are ready to use a stress management system if it is integrated into the university online platform. This gives evidence that undergraduates wish to follow a stress management system to help them relieve their stress. Out of 73 responses, 54.8% of the respondents wish to use the system very likely while 31.5% of the respondents also wish to use the system.

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.
- Student should be able to chat with Counsellor of the University.
- Also, Student should be able to schedule a meeting by clicking on Schedule.
- Admin should be able to add activities with regard to particular stressor by accessing the different module 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 counsellors, 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 mentioned coping strategies and stressors in the interview by counsellors. Another mentioned point in the interview was the student may have 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 be auditable. (Logs)
- The system should be able to respond promptly.
- The system should be able to access by many students at once.

4) Justification of Non-Functional Requirements

Data Privacy and Confidentiality of a student was their main concern of them. As mentioned above, security methods like two factor authentication and multi factor authentication are being used. And role-based permission is given by the system for all the modules. Since this system is used by students who struggle with stress, the User Interface is designed to give more user experience and it is easy to use.

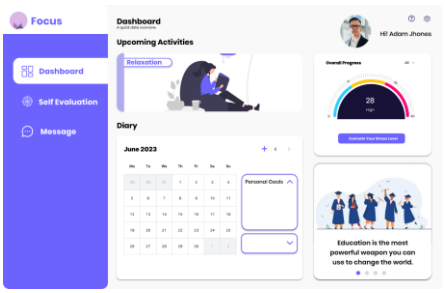


Fig 1. Student Dashboard

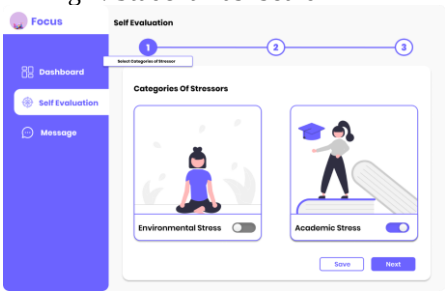
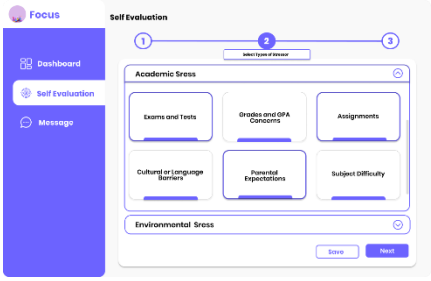


Fig 2. Self-Evaluation (Step 01

Fig 3. Self-Evaluation (Step 02)

C. Result on Design Evaluation

Once UI design was completed, the prototype was shown to the selected students and one counsellor. The students were satisfied regarding self-evaluation and stress measurement features that were going to be implemented in the system. The counsellor was also satisfied but she gave some more opinions regarding the system. The one concern of her is to change the theme colour of the system as per the desire of the student. So, this option was added in the additional module called set up and admin can give permission to permissions to change colour. So basically, students can select the primary colour and the secondary colour. According to their preference the colours of the systems can be changed. And chat option and option to schedule a meeting was satisfied. Since she mentioned during her interview, this app may not be a problem for existing counselling service in the university. Also, another satisfying point of her was this option helps to university to know about the counselling service in the university and students can get engaged with the counsellor through this option. There are some suggestions from students to give options for time management, to make priority lists and many more. So, these features will be taken to the second phase of this project.



So, they were satisfied with the design system, and they think this helps to reduce stress of the students by using this app. Both positive and negative opinions taken into the

account, and this will be implemented in the second phase of this project.

IV. Conclusion

First, deeper understanding of the subject was given throughout the research project, particularly in relation to stress management among undergraduates. This knowledge has improved my academic performance as well as my ability to meaningfully participate in conversations about mental health and wellbeing. The research process also helped me develop my analytical and critical thinking abilities. My capacity to properly synthesize knowledge has improved because of reading various literature reviews and materials. Also, throughout this research project, I have learned how to manage time and balance workload. By managing time and balancing workload, I have been able to prioritize tasks and improve my strengths.

Throughout this research project, there have been rewarding and challenging learning curves. One significant learning curve is reading literature reviews and analysing them to find a research gap well. Another significant learning curve is Data Gathering. From designing questionnaires to conducting interviews, I gained practical experience of research methodology. I also improved my ability to interpret and analyse data. Also, I have learned about conducting interviews since two interviews were conducted with counsellors.

V. Future Works

With the limited time, the system cannot be implemented as per the design. So, the system will be implemented according to the designed UIs in Chapter 5 and above-mentioned technologies in Chapter 4 under 4.6 section. Also, according to the requirements of universities new features will be added further and to increase more accuracy of measurement of stress level and suggestions of coping strategies AI, IOT devices can be integrated.

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] Fernando, M., Chandrasiri, A., Dayabandara, M., & Reavley, N. J. (2021). Cultural adaptation of mental health first aid guidelines for depression for Sri Lanka: a Delphi expert consensus study. *BMC Psychiatry*, *21*(1). https://doi.org/10.1186/s12888-021-03598-0

[5] 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

[6] 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

[7] Gao, W., Ping, S., & Liu, X. (2020). Gender differences in depression, anxiety, and stress among college students: A longitudinal study from

China. *Journal of Affective Disorders*, *263*, 292–300. https://doi.org/10.1016/j.jad.2019.11.121

[8] 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

[9] 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

[10] 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

[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] Karaman, M. A., Lerma, E., Vela, J. C., & Watson, J. C. (2019a). Predictors of Academic Stress Among College Students. *Journal of College Counseling*, *22*(1), 41–55. https://doi.org/10.1002/jocc.12113

[14] 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

[15] 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

[16] 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

Emoland: Enhancing Facial Emotion Skills Among Children with Autism Spectrum Disorder (ASD) in Sri Lanka

Devanga Kavikeshawa
Faculty of Computing
NSBM Green University
Colombo, Sri Lanka
chandunidevanga@gmail.com

Rasika Ranaweera
Faculty of Postgraduate Studies & Professional Advancement
NSBM Green University
Colombo, Sri Lanka
ranaweera.r@nsbm.ac.lk

Abstract*—Autism Spectrum Disorder (ASD) is a neurodevelopmental disorder characterized by diverse symptoms, including challenges in emotional intelligence. Children with ASD often struggle with recognizing and expressing facial emotions, significantly affecting their overall well-being. Currently, there is a lack of accessible and effective early intervention tools tailored for children with ASD in Sri Lanka, specifically within the critical age range of 2-7 years. This research aims to bridge this gap by leveraging Information Communication Technology (ICT) and Artificial Intelligence (AI) advancements to develop a user-friendly mobile application, functioning as an early intervention tool with a focus on facial emotion recognition and production skills. The proposed solution incorporates Ekman's six basic emotions methodology for facial emotion recognition, progressive learning materials featuring familiar individuals to children with ASD, and real-life context-based evaluations to monitor children's progress. After conducting a comprehensive analysis and gathering requirements, a prototype solution, including all the required features, was developed. The findings and the analysis of solution effectiveness through evaluation strongly indicate that the system has significant potential.

***Keywords*—Autism Spectrum Disorder, Facial Emotion Recognition, Sri Lanka, Early Intervention, Computer Vision, DeepFace, Mobile Application**

I. INTRODUCTION

Autism Spectrum Disorder (ASD) includes a spectrum of neurodevelopmental challenges, involving deficits in social skills, communication, and cognitive functions such as facial emotion recognition and Theory of Mind (ToM). Individuals with ASD frequently demonstrate lower Emotional Intelligence (EI), especially in interpreting emotions from the environment, connecting emotions to others' experiences, comprehending emotions from contextual cues, and regulating one's own and others' emotions [1].

Facial emotion recognition is a critical aspect of non-verbal communication and social interaction, significantly impacting social interactions, effective communication, and the overall well-being of individuals with ASD [2] [3]. Recognizing this, it becomes evident that interventions addressing these challenges are essential for enhancing the quality of life for children with ASD, especially within the critical age range of 2-7, where cognitive and emotional development is particularly significant.

The ability model of EI, as demonstrated by Ekman's six basic emotions, offers a conceptual framework for comprehending emotional deficiencies in individuals with ASD. Empathy, a key component of EI, involves conveying one's mental states to others and responding appropriately. Individuals with ASD often exhibit innate or heightened systemizing abilities, suggesting

the potential for improved cognitive empathy, particularly in recognizing emotions [4]. Understanding and managing emotions are critical for children's development, facilitating effective communication and navigation of their social environment.

Despite the global recognition of challenges faced by individuals with ASD, the situation is particularly challenging in Sri Lanka. Families with autistic children encounter difficulties in various aspects of life, including housework, finances, parental emotional and mental health, and a limited ability to respond to the needs of other children within the family. Sri Lankan schools often discourage the enrolment of students with ASD due to a lack of understanding and resources, leading to additional obstacles for families, such as stigma, limited employment prospects, and safety concerns.

The term "early intervention" is pivotal in addressing the developmental needs of children with ASD under the age of seven. Early intervention, crucial for enhancing the quality of life for autistic children, involves play-based interventions targeting critical brain developmental stages [5]. Emphasis is placed on the importance of early intervention to mitigate the negative effects of neurodevelopmental issues, as recommended by the American Academy of Pediatrics for developmental and behavioral screenings at 9, 18, and 30 months.

However, existing interventions encounter challenges related to shared difficulties among therapists and parents, unique challenges in facial emotion recognition specific to ASD, an emphasis on language and behavior abilities, and a lack of accessible and effective Information and Communication Technologies (ICT)-based interventions. ICTs play a significant role in addressing these challenges, offering tools for communication, emotional exploration, and skill development in individuals with ASD [3]. Notably, mobile computing has facilitated the integration of Facial Emotion Recognition (FER) systems into smartphones, rendering it a viable platform for enhancing emotional intelligence applications for children with ASD [6].

While facial emotion training has demonstrated effectiveness in neurotypical adults, its impact on children with ASD, particularly in the Sri Lankan context, remains understudied. This research aims to bridge this gap by developing a user-friendly mobile application that leverages artificial intelligence to enhance facial emotion recognition and production skills in children with ASD aged 2-7 in Sri Lanka. By addressing these challenges through an ICT-based intervention, this research seeks to make a significant contribution to the existing body of knowledge and provide a practical solution for the unique needs of children with ASD in Sri Lanka.

II. Literature Review

This research includes a comprehensive analysis of existing systems to identify research gaps. The investigation of various studies has yielded insights into the algorithms, design, and workflows employed in similar systems. Addressing these identified research gaps is crucial for advancing the development of effective solutions that specifically target the enhancement of facial emotion recognition and production skills in children with ASD.

A. Existing Systems Analysis

1) Facial Expression Wonderland (FEW) [2]

FEW is an innovative computer game design prototype developed with the primary goal of enhancing the recognition of facial emotions in children with ASD. Inspired by the Walt Disney movie "Alice In Wonderland," FEW utilizes animated characters and voice prompts to address ToM through diverse emotional situations. While it offers multiple training levels, potential concerns include the appropriateness of cartoon expressions for emotion testing, the need for careful emotion selection, and the effectiveness of difficulty progression. Real experiments and eye-tracking are deemed necessary for accurate gaze detection during gameplay.

2) The Serious Game JEMImE [7]

JEMImE is a serious game designed for children with ASD to improve facial expression production in social contexts. The game comprises training and playing phases, with training involving imitation and production of facial emotions accompanied by real-time feedback. Developed in C# using Unity 3D, it utilizes Pairwise Conditional Random Forests for facial expression assessment and Curapy.com for data recording. Despite yielding promising results, the study acknowledges limitations, including a small sample size, a one-hour game duration, and the need for a child-specific facial expression classifier.

3) The "Let's Face It!" Program [8]

"Let's Face It!" utilizes a comprehensive psychological battery with seven interactive computer games, addressing face processing deficits in ASD. These games are categorized into three domains: attention to faces, recognition of facial identity and expression, and interpretation of facial cues in social contexts. The inclusion of engaging graphics, music tracks, and progressively challenging levels serves to sustain children's interest and motivation. While the program has demonstrated success in teaching emotional skills, challenges persist, including variations in effectiveness, uncertain generalizability to real-life situations, and potential difficulties in transferring skills beyond computerized contexts.

4) The Transporters DVD Series

This is designed for ages 2-8, utilizes animated vehicles to depict emotions. A study assessing the series demonstrated substantial enhancements in emotion comprehension and recognition skills among children with ASD [4]. Despite its effectiveness in improving emotion comprehension, the DVD format lacks the accessibility and user-friendliness typically associated with mobile applications.

5) EmoTEA [9]

This is a serious game developed for Android, targeting ASD children aged 6 to 12. It integrates Ekman's six-basic-emotions methodology, Affective Computing, and Tangible User Interfaces (TUIs) to augment emotional intelligence. The observed positive outcomes in emotion recognition validate its effectiveness. However, limitations include the specificity of the age range and potential disruptions to routines, underscoring the importance of careful implementation and individualized approaches.

6) World of Kids [10]

This is a personalized mobile and tablet game series, utilized facial emotion recognition algorithms and voice recognition tools for customized learning. The application features carefully designed user interfaces and graphic cues to enhance comprehension. The games are created using AndEngine, a 2D game engine library. Although initial tests demonstrate high efficiency, limitations include dependence on a local SQLite database, affecting storage capacity and accessibility.

7) Emotions on the Go [11]

This introduces a mobile-based emotion detection system utilizing OpenFace for real-time facial expression assessment. The researchers used technologies such as Orbeus ReKognition API, OpenFace, Fraunhofer SHORE, and the AFFDEX SDK. The study highlights OpenFace's suitability for Android but acknowledges limitations, including a one-week training duration and a small initial sample size, emphasizing the need for further research.

B. Technological Analysis

1) Machine Learning Algorithms

Recent literature highlights the significance of machine learning algorithms. Support Vector Machines (SVMs) have demonstrated effectiveness in categorizing high-dimensional facial expression data [5] [12]. The Viola-Jones Feature Detection System utilizes rapid object detection and grayscale image enhancement for real-time emotion recognition, effectively addressing luminosity variations [10]. Additionally, a modified eyemap and mouthmap algorithm, implementing TensorFlow for neural network execution, has shown efficiency in

facial geometry detection and emotion classification [13].

2) Deep Learning Algorithms

In deep learning algorithms for facial emotion recognition, Convolutional Neural Networks (CNNs) emerge as a prevalent choice due to their automatic learning capabilities, particularly in computer vision tasks. The effectiveness of CNN architectures is demonstrated through the utilization of techniques such as ReLU activation, maximum pooling, and dropout to enhance accuracy [5]. Additionally, pre-trained models, including ResNet50, MobileNetV2, and VGG19, show notable success in infant facial emotion recognition. Transfer learning emerges as an efficient approach by leveraging pre-trained models and fine-tuning for facial emotion recognition [14]. Google Teachable Machine provides a user-friendly avenue for creating custom classification models through transfer learning, facilitating model integration into lightweight web apps for diverse applications [5] [15].

3) Computer Vision Libraries

OpenCV is an open-source tool for image processing across various platforms, playing a vital role in machine learning and deep learning for training and evaluating FER models [5].

Notably, the DeepFace framework, developed by Emre Serengil, has garnered attention for its remarkable accuracy in face identification tasks. Leveraging CNNs, DeepFace excels in facial attribute analysis, encompassing age, gender, emotion, and race. The framework integrates pre-trained state-of-the-art models such as VGG-Face, Google FaceNet, OpenFace, Facebook DeepFace, DeepID, ArcFace, Dlib, and SFace, showcasing its versatility [16]. An analysis of DeepFace's CNN architecture reveals simplicity and effectiveness, achieving competitive accuracy in emotion prediction when evaluated against benchmark datasets [17]. The combination of OpenCV and the DeepFace framework emerges as a powerful toolset for facial emotion analysis within various domains, including edge computing and educational technology [18].

C. Reflection

The comprehensive analysis of existing literature highlights a significant research gap concerning the long-term impacts of ICT-based interventions on facial emotion recognition skills in children with ASD. Existing studies have less focus on the crucial aspect of social skills, particularly recognizing and expressing emotions and the unique challenges in emotion recognition. The technological analysis further reveals limitations in the application of CNNs, particularly in dealing with one-dimensional training data, challenges in training stability, and feature duplication. Moreover, the lack of interventions in the Sri Lankan context emphasizes the urgency to address this gap. In light of these challenges, there is a compelling need to bridge the identified gap, particularly within the Sri Lankan context, offering a tailored and accessible solution for the targeted age group.

III. Methodology

A. Data Collection

In this research, a mixed-methods approach will be applied, combining both quantitative and qualitative methodologies. The data collection mechanisms include interviews, expert opinions, and questionnaires for both pre-analysis and post-analysis phases. Comprehensive interviews with therapists, special needs education teachers, doctors, and parents during the pre-analysis phase aim to differentiate perspectives and existing interventions for children with ASD, revealing crucial insights to identify gaps and requirements for the proposed solution. In the post-analysis phase, discussions with experts provide valuable feedback on the implemented application, assessing its effectiveness, uncovering strengths and weaknesses, and guiding future design improvements. Additionally, a post-analysis questionnaire administered to parents through Google Forms facilitates the

evaluation of the solution's effectiveness through statistical techniques.

B. Data Analysis

Thematic analysis was used to analyze interview data from stakeholders, including Ms. Jayani Panapitiya, Mrs. S. D. Rathuwithana, therapists Mrs. Chamila and Mr. Bharatha from Cherish Foundation, and Dr. Savithri Dias. Identified themes highlighted the importance of visualization, colour preferences, multisensory learning, repetitive teaching, current methods, challenges, technical proficiency, and recommended features. Stakeholders highlighted current challenges such as children's reliance on parents' facial expressions, negative home experiences impacting school, and acknowledged technical proficiency in using smartphones with the need for parental guidance in technology use. Key recommendations included a minimalist user interface, familiar images, progress tracking, and engaging materials. These findings form the basis for developing a user-friendly mobile application that leverages AI to enhance facial emotion recognition and production skills in children with ASD aged 2-7, addressing critical gaps identified in existing literature and methodologies.

IV. Solution

A. Proposed Solution

The proposed solution, named "EmoLand," presents a high-level conceptual framework addressing the identified research gap by leveraging Ekman's six basic emotions methodology for facial emotion recognition in children with ASD. Based on the findings, a mobile application was considered the most suitable choice for development. The solution incorporates learning materials featuring familiar individuals, primarily parents, to enhance relatability for children. It adopts a progressive learning approach, integrating lessons, real-life social context evaluations, and progress tracking to improve comprehension and expression of facial emotions. Children actively participate by scanning their faces for given emotions, enabling facial expression analysis to track learning progress. Motivated by outcomes, the solution aims to effectively support and enhance the learning experience, providing a tailored ICT-based intervention for the development of facial emotion skills in children with ASD aged 2-7.

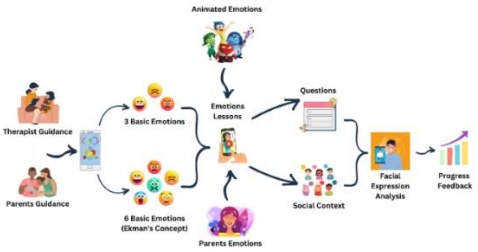


Fig. 1. High level graphical representation of the conceptual solution

B. Implementation

1) Technology

Python was used to construct the facial emotion recognition model, utilizing popular libraries such as OpenCV with Haar Cascade classifier and NumPy for image processing and face detection. The DeepFace framework played a crucial role in facial attribute analysis. Flask was used for API development, and Ngrok securely hosted the Python APIs. Flutter framework was used to develop a visually appealing and user-friendly mobile application. The Android Emulator within Android Studio was utilized for testing and running the Flutter app. Simultaneously, the React framework was used for the web-based frontend, with a specific emphasis on real-time facial emotion detection. Firebase Cloud Firestore served as a NoSQL database for storing user-related information. The system architecture is displayed in Fig 2.

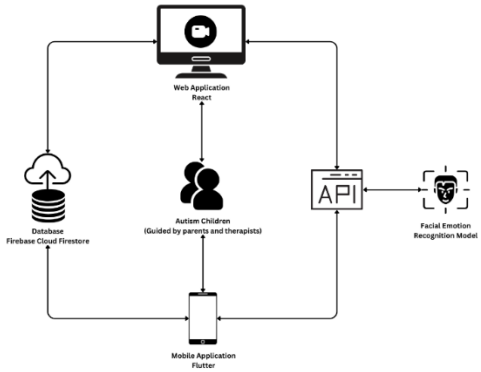


Fig. 2. System Architecture Diagram

2) Facial Emotion Recognition Model

FER model is an essential component of the system, designed to assess the facial emotion production skills of children. As illustrated in the Fig 3, after the completion of emotion lessons in the mobile application, the model assesses the child's ability to express emotions, evaluating the accuracy of their expressions and application in real-life social scenarios. The outcomes generated by the FER model are utilized to provide feedback within the application interface, delivering valuable insights to enhance the child's emotional development.

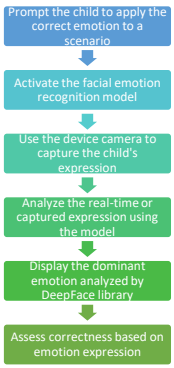


Fig. 3. Process of Facial Emotion Recognition Model

FER model is implemented through two methods. Firstly, for image-based emotion recognition, Python code utilizes the DeepFace library for emotion analysis and OpenCV for face detection. The code processes an input image, converts it to RGB format, identifies and marks faces using a Haar Cascade classifier, and overlays the dominant emotion on the image. Secondly, for real-time emotion recognition in video streams, the Python code captures frames from a webcam, applies face detection using OpenCV, and utilizes DeepFace for emotion analysis. The live feedback of the dominant emotion is superimposed on the video feed, providing real-time emotion recognition. Both models are converted into Flask APIs, facilitating communication between the FER model and the frontend application, ensuring seamless integration into mobile and web-based environments.

3) Designing User Interfaces

The user interface (UI) design for the 'EmoLand' mobile application prioritizes creating a practical and user-friendly experience tailored to the mindset of ASD children. Recognizing the importance of predictability and consistency for ASD children, the design carefully considers factors. Literature emphasizes the role of color psychology in influencing emotions and behaviors in ASD individuals, with personal experiences and cultural factors influencing these preferences. The UI integrates easily readable fonts and minimizes complex elements to enhance usability, considering the reading comprehension challenges often faced by individuals with ASD. Moreover, the design incorporates graphic cues, such as images and icons, to illustrate concepts and improve reading comprehension, aligning with findings from various studies and data analyses. These design considerations aim to create an inclusive and supportive environment for children with ASD.

4) Main Features

The "EmoLand" mobile application features several carefully selected components rooted in findings from existing research and interview data analysis. Leveraging familiar images of parents as a potent teaching tool, the app facilitates facial emotion learning through a unique parent emotions feature. Customized lessons focus on three fundamental emotions, utilizing parent emotion images or engaging cartoon depictions for a comprehensive learning experience. The application incorporates an evaluation component, where children

answer questions related to emotions, express corresponding emotions on their faces, and submit images for assessment using facial emotion recognition. Progress tracking is enabled through a comprehensive database, offering insights into a child's development. A streamlined registration process ensures a personalized learning experience for children in alignment with the proposed solution's objectives.

V. Evaluation and Results

A. Model Accuracy Testing

The accuracy of the developed FER model, utilizing the DeepFace library, was evaluated through a test dataset including sample images representing seven emotions. The analysis, facilitated by Python libraries revealed varying accuracies across different emotions. Notably, happy emotion detection exhibited the highest accuracy, while disgust emotion encountered challenges with no accurate predictions. The comparison highlighted insights into the confusion matrix, indicating common misclassifications between certain emotions.

In parallel, the accuracy of the implemented model was benchmarked against a model constructed and trained using Google Teachable Machine. The Teachable Machine model, trained on a hybrid dataset, demonstrated varying accuracy across emotion classes. Notably, while the model performed well in identifying happy expressions, it exhibited lower accuracy in classifying angry and sad facial expressions. The comparison served to validate the accuracy of the custom model developed using the DeepFace library, justifying its selection for the FER system. The analysis of both models provided valuable insights into their performance, focusing on areas of strength and potential improvement, particularly in handling certain emotions.

B. Solution Effectiveness Analysis

1) Intervention Results

The prototype of the "EmoLand" application was tested with three ASD children (Child 1, Child 2, Child 3) currently undergoing therapy and experiencing mild struggles in emotional skills. They completed all six lessons within the application. The evaluation results indicate varying success levels among the children as mentioned in the Table I.

TABLE I. INTERVENTION RESULTS

Child	Age	Lessons Status	Evaluation Status		
			Happy	***Sad***	***Angry***
1	7	All Completed	Pass	Pass	Fail
2	7	All Completed	Pass	Fail	Fail
3	6	All Completed	Pass	Pass	Pass

These findings collectively suggest that the developed application prototype has varying degrees of effectiveness for different children. While Child 3 exhibited success in recognizing and producing all three emotions, Childs 1 and 2 faced challenges with specific emotions. These results suggest that while the application effectively supports some children in recognizing and producing facial emotions, individual differences and challenges influence the outcomes. Further insights and refinements may be necessary to tailor the application to the unique needs of each child, ensuring a more personalized and adaptive learning experience.

2) Expert Opinions and Survey Results

The developed "EmoLand" mobile application successfully achieves its main objective of serving as a user-friendly AI-based early intervention tool for children with ASD aged 2-7 in Sri Lanka. Expert discussions with domain experts, as well as a satisfaction survey with parents, confirm the positive impact and effectiveness of the application in aiding the emotional development of ASD children. The application, designed based on identified preferences and challenges of ASD children, received commendation for its visual learning approach, integration of familiar images, and utilization of ICT-based intervention. Additionally, the survey results

indicate high levels of satisfaction among parents, emphasizing the application's utility and effectiveness. The objectives related to providing a useful tool for parents and therapists, analyzing the effectiveness of familiar images, evaluating ICT-based intervention, and examining real-life application of learned emotions are all successfully met, as evidenced by expert feedback and survey responses. The study concludes with valuable suggestions for further improvements, ensuring ongoing enhancement and adaptability to the evolving needs of ASD children.

VI. Ethical Considerations

The process of gathering personal data was impacted due to ethical concerns. It was not feasible to categorize data based on age, as the study pertained to children with ASD, and collecting data related to their ages was not practical. Privacy concerns affected in sharing personal data and engaging in direct communication. Furthermore, it is important to acknowledge that ethical considerations and time constraints limited the sample size to only three children for testing the application, which may restrict the comprehensive evaluation of system effectiveness. Additionally, given the need for repetitive learning in ASD children, the relatively short testing duration of approximately 2-3 hours highlights the necessity for further evaluation to thoroughly assess the application's impact.

VII. Future Recommendations

The "EmoLand" application demonstrates promising potential for future enhancements. Recommendations include expanding the emotion list within the system to include a broader range, diversifying lesson delivery methods, and introducing more comprehensive evaluation techniques. To enhance personalization, the system could tailor content based on individual needs, considering factors like age, cognitive abilities, and learning styles. Integrating voice instructions and optimizing the user interface for interactivity would foster a more engaging experience. Furthermore, improving the FER model's accuracy is crucial, suggesting the exploration of neural network development for improved accuracy and effectiveness in emotion detection and analysis. These proposed enhancements aim to create a more adaptive, engaging, and effective platform, addressing the unique needs and diverse learning styles of children with ASD.

VIII. Conclusion

The primary objective of this research project is to investigate and develop an ICT-based solution aimed at enhancing the facial emotion recognition and expression skills of children with ASD in Sri Lanka. Through a comprehensive analysis and the collation of requirements, a prototype solution was developed, incorporating the necessary features. The findings and effectiveness analysis strongly suggest that the system holds significant potential. Given the current state of development and the challenges faced, the application is poised for future improvements.

Acknowledgment

I express gratitude to Dr. Rasika Ranaweera, my supervisor and Dean of the Faculty of Postgraduate Studies at NSBM Green University Town, for invaluable guidance. Special thanks to Special Needs Education Professionals, Therapists, Doctors, and Parents for sharing their insights, enriching this study. I appreciate Dr. Mohamed Shafraz and Faculty of Computing for their support. The unwavering encouragement from my family and friends has been a constant motivation. While it's challenging to individually acknowledge all contributors, I extend heartfelt thanks to everyone for their collective efforts in making this research project a success.

References

[1] D. A. Trevisan, E. A. Abel, M. A. Brackett, and J. C. McPartland, "Considerations About How Emotional Intelligence can be Enhanced in Children With Autism Spectrum Disorder," Frontiers in Education, vol. 6, Apr. 2021, doi: 10.3389/feduc.2021.639736.

[2] R.-Y. Tseng and E. Y.-L. Do, "Facial expression wonderland (FEW)," *Proceedings of the 1st ACM International Health Informatics Symposium,*

Nov. 2010, doi: https://doi.org/10.1145/1882992.1883064.

[3] E. Papageorgiou, "The Impact of Information and Communication Technologies on the Emotional Education of Children with Autism Spectrum Disorder. Views of Specialist Educators and Psychologists," *OAlib*, vol. 07, no. 05, pp. 1–24, Jan. 2020, doi: 10.4236/oalib.1106248.

[4] S. Baron-Cohen, O. Golan, and E. Ashwin, "Can emotion recognition be taught to children with autism spectrum conditions?," *Philosophical Transactions of the Royal Society B*, vol. 364, no. 1535, pp. 3567–3574, Dec. 2009, doi: 10.1098/rstb.2009.0191.

[5] M.-L. Gau *et al.*, "Effectiveness of using artificial intelligence for early child development screening," *Green Intelligent Systems and Applications*, vol. 3, no. 1, pp. 1–13, May 2023, doi: 10.53623/gisa.v3i1.229.Al-Saadi and D. Al-Thani, "Mobile Application to identify and recognize emotions for children with autism: A systematic review," *Frontiers in Child and Adolescent Psychiatry*, vol. 2, Mar. 2023, doi: 10.3389/frcha.2023.1118665.Grossard *et al.*, "Teaching Facial Expression Production in Autism: The Serious Game JEMIME," *Creative Education*, vol. 10, no. 11, pp. 2347–2366, Jan. 2019, doi: 10.4236/ce.2019.1011167.

[6] J. W. Tanaka *et al.*, "Using computerized games to teach face recognition skills to children with autism spectrum disorder: the Let's Face It! program," *Journal of Child Psychology and Psychiatry*, vol. 51, no. 8, pp. 944–952, Jul. 2010, doi: 10.1111/j.1469-7610.2010.02258.x.

[7] J. M. Garcia-Garcia, M. del M. Cabañero, V. M. R. Penichet, and M. D. Lozano, "EmoTEA," *Proceedings of the XX International Conference on Human Computer Interaction*, Jun. 2019, doi: https://doi.org/10.1145/3335595.3335639.

[8] N. Heni and H. Hamam, "Design of emotional educational system mobile games for autistic children," *2016 2nd International Conference on Advanced Technologies for Signal and Image Processing (ATSIP)*, Mar. 2016, doi: https://doi.org/10.1109/atsip.2016.7523168.

[9] T. Kosch, M. Hassib, R. Reutter, and F. Alt, "Emotions on the Go," *International Conference on Advanced Visual Interfaces (AVI '20)*, Sep. 2020, doi: 10.1145/3399715.3399928.

[10] S. B. Punuri *et al.*, "Efficient Net-XGBOOST: an implementation for facial emotion recognition using transfer learning," *Mathematics*, vol. 11, no. 3, p. 776, Feb. 2023, doi: 10.3390/math11030776.

[11] Joseph and P. Geetha, "Facial emotion detection using modified eyemap–mouthmap algorithm on an enhanced image and classification with tensorflow," *The Visual Computer*, vol. 36, no. 3, pp. 529–539, Feb. 2019, doi: 10.1007/s00371-019-01628-3.

[12] M. a. H. Akhand, S. Roy, N. Siddique, A. S. Kamal, and T. Shimamura, "Facial emotion recognition using transfer learning in the deep CNN," *Electronics*, vol. 10, no. 9, p. 1036, Apr. 2021, doi: 10.3390/electronics10091036.

[13] Y. P. Pujari, D. Prasad, N. Malleswari, and N. Gupta, "Implementation of Machine Learning Based Google Teachable Machine in Early Childhood Education," *International Journal of Early Childhood Special Education*, vol. 14, no. 3, pp. 4132–4138, May 2022.

[14] Serengil, "GitHub - serengil/deepface: A Lightweight Face Recognition and Facial Attribute Analysis (Age, Gender, Emotion and Race) Library for Python," *GitHub*. https://github.com/serengil/deepface

[15] Atanassov and Dimitar Pilev, "Pre-trained Deep Learning Models for Facial Emotions Recognition," Oct. 2020, doi: https://doi.org/10.1109/icai50593.2020.9311334.

[16] T. RAJKUMAR, INDUMATHI, and SANTHANA MEGALA, "Novel learning management system (LMS) with Deepface Emotion Analysis," *Advances and Applications in Mathematical Sciences*, vol. 21, no. 9, pp. 5053–5061, Jul. 2022, doi: 10.1109/icai50593.2020.9311334.

Web Application for Patients Providing Predictions and Understandable Summarized Information of Medical Test Report Results via the Machine Learning Algorithm

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

Dr. A C V Silva
Department of Neuro Anaesthesia)
National Hospital of Sri Lanka
Colombo, Sri Lanka
charakavsilva@gmail.com

Abstract*— Healthcare domain is an important domain which exists in society. In Healthcare domain health literacy is a major factor. Health literacy among 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 with 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. 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. In this work a web application is developed as the solution to the problem which is mentioned above with React Js and Python technologies.

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. This might lead to risky outcomes.

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 uses same concept of the suggested system. For this purpose, different papers have 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*[3]

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. 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* [4]

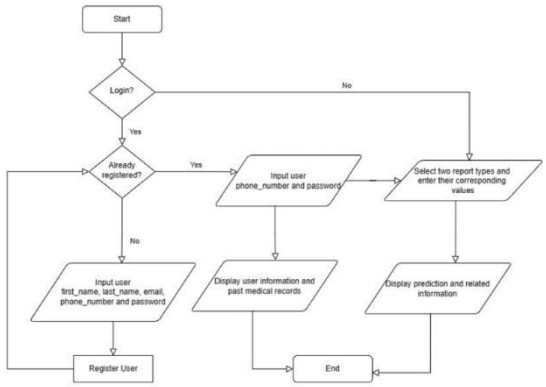
Predicting the survival rate of Titanic disaster using machine learning approaches discusses about the model suggested to predict the survival rate of titanic disaster using machine learning approaches. Four algorithms are implemented to the model, which are e Decision Tree, Logistic Regression, Naïve Bayes, and Random Forest algorithms. 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*[5]

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.

D. A Machine Learning Approach for Student Assessment in E-Learning using Quinlan's C4.5, Naïve Bayes, and Random Forest Algorithm.[6]

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 Decision Trees-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. As



the result proposed 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. METHOD

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% based on an evaluation conducted which showed the highest accuracy.

B.

C. Proposed Model and workflow of the proposed system.

In the proposed system, 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 directs to a webpage that gives more information about the reports that the user has mentioned. 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.

D. 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.

Figure 1. Proposed System in a flowchart

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 is communicated using APIs.

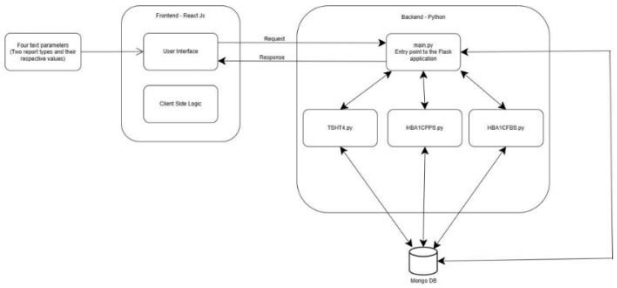


Figure 2. System workflow of the developed web

IV. Evaluation and Results

Algorithm for the models was chosen by evaluating four machine learning algorithms 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.

1) Training data : Testing data – 80% : 20%

Table 2.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%

2) Training data: Testing data – 75%: 25%

Table 3.. Evaluating 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%

3) Training data: Testing data – 70%: 30%

Table 4.. Evaluating of algorithms based on the ratio training data: testing data - 75%: 25%

	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 analysing 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 was as expected, and the requirements of the systems were achieved.

V. 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 analysed 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.

VI. Conclution

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. 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.

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 Thilakarayhne, for providing me with unwavering support, guidance and expertise thought out this research project. All his insightful advice and encouragement have been invaluable. 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?," Journal of Health Communication, vol. 15, no. SUPPL. 2. pp. 9–19, 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] 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.

[4] 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.

[5] 2018 Fourth International Conference on Computing Communication Control and Automation (ICCUBEA).

[6] 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.

Students' Higher Education Performance Prediction and Monitoring via Machine Learning

Nonimi Bandara
Dept. of Computer and Data Science
National School of Business Management
Homagama, Sri Lanka
nonimitharuni@gmail.com

Gayan Perera
Dept. of Computer and Data Science
National School of Business Management
Homagama, Sri Lanka
gayanp@nsbm.ac.lk

Abstract*—The prediction of student performance, which involves using data analytic techniques to predict the likelihood that students will perform well on academics, is a crucial aspect of the research on education sector in order to create effective educational strategies and measures it helps to identify the students' potential strengths and weaknesses. Here machine learning techniques play a huge role in predicting students' performance efficiently and effectively. In this study, a machine learning solution will be implemented using Naïve Bayes classification to predict students' performance at the GCE A/L examination aiming reducing the high failure rate of Physical Science stream. Parallelly, important statistical data of higher education sector will be visualize using an online platform for the different beneficiaries who may interested in higher education sector.

Keywords*—Education domain, Students' performance prediction, physical science stream, GCE A/L, Machine learning, Naïve Bayes classification, Data visualization

I. INTRODUCTION

Education is the leading potential and success of individuals and societies also it is a human right in every nation. It grants people the knowledge, abilities, and critical thinking skills they need to cope with their challenges in life and make a meaningful contribution to their communities. People gain knowledge of the world, its cultures, histories, and scientific developments through education. It encourages social and economic mobility, giving people the support they need to escape the cycle of poverty and enhance the quality of their lives.

In Sri Lanka every child is born with the right for the free education up to university education and it divide into three stages as primary education, junior secondary education and senior secondary education. At senior secondary education firstly, students have to face for the General Certificate of Education (Ordinary Level-G.C.E.(O/L)) at grade 11 and at grade 13 the General Certificate of Education (Advanced Level-G.C.E.(A/L)) is held as the university entrance exam [1]. These major exams are conducted by the Ministry of Education in Sri Lanka [10].

At G.C.E.(A/L) here students make a critical decision since A/L subjects stream determines the future carrier path of the majority students. Due to economic and social reasons students tend to choose physical science stream since they believe it will support them to settle with a financially stable occupation which will leads to skill less workforce [11]. Also, it affects to high failure rate in physical science stream [13].

So, the main objective of this study is to address reducing high failure rate of physical science stream by predicting students' performance at GCE A/L examination using machine learning to support students who needs improvements [12]. Also, higher education statistical data visualization on a web platform to provide important information such as graduates, undergraduates count per each year, number of students who are enrolled in higher education by government or non-government educational institutions etc. will be

developed for the beneficiaries who are interested on higher education sector.

II. LITERATURE REVIEW

The studies on education domain and existing systems that have been introduced to predict the student performance at examinations and technologies they have used to define solutions were discussed using previous literatures. Those approaches and their uses in current systems have been investigated and the suitability of those approaches has been considered by analyzing supporting literatures.

A. *Subject Stream Prediction: A Machine learning Approach to Select the Suitable Subject Stream for Senior Secondary Students in Sri Lanka[2]*

In this study, a system is proposed to predict the suitable A/L subject stream for students by considering their previous academic achievements up to GCE O/L examination. This system was developed using 1000 employees' data and it support the student to choose their A/L subject stream as their skills and preferred carrier paths. This system is developed with high accuracy (72) using Random Forest Algorithms.

B. *Deep Learning with Data Transformation and Factor Analysis for Student Performance Prediction[3]*

This study presents an approach to build a student performance prediction model in the next semester based on the previous semester achievements using Convolutional Neural Networks and Long Short-Term Memory techniques. Also, they have analyzed several techniques for data preprocessing. Their results were feasible for applying practically.

C. *Prediction of Student's Performance Using Machine Learning[4]*

The introduced system in this paper has implemented as a novice machine learning tool to predict the future score and accordingly help students to excel at every stage. It was focused on excelling at every course for overall academic success. This model will keep the track and overall growth of every student.

D. *Prediction of Student Academic Performance using Neural Network, Linear Regression and Support Vector Regression: A Case Study[5]*

This is a case study that was designed to support higher education institutions to predict their students' performance level at the early stages to decrease dropouts used Neural Network, Linear Regression and Support Vector Regression for predicting GPA without considering other economic, social, or psychological factors. In this study the results obtained from the separate technologies were evaluated.

III. METHODOLOGY

While preparing for the examination many students fall behind and this is often in physical science stream. If there is a way to support them to identify their level while preparing for the exam, it will improve the overall students' performance at their final examination and gradually it will reduce the failure rate of physical science stream. Therefor this study will be conducted to predict students' performance at the final examination and online platform will be implemented to visualize statistical data to give a proper insight on higher education sector.

A. Data Acquisition Stage

The data collection mechanism for this research project primarily employed a

quantitative approach. Data were gathered from a small-scale education institute through systematic extraction of quantitative information from their records. Academic records on past three term tests, attendance records, and demographic information such as gender, parents education level were all included as variables. After collecting required amount of data the dataset was carefully constructed including all required data to predict the performance with considerable accuracy and to protect the integrity.

B. The proposed system

The proposed system predicts the student's performance using machine learning technologies [15] by accepting three previous term test results of Combined Mathematics, Physics and Chemistry subjects which are the main subjects of physical science stream with other external factors such as attendance, parent education and extra classes. After student enter their details, system will predict the performance accordingly. So that students can identify their performance level and improve as needed or seek support. Also, for the data visualization about higher education sector, important statistical data such as graduates, undergraduates count per each year, university students count for each stream, student enrollment in higher education will be represented by charts for better presentation. This will be a support for the students, educators, researchers and any other party that involves in higher education for their purposes.

C. Implementation

The solution is implemented as a web platform since it is easy to accessible for any user and can use without any installation. The website is developed with HTML and CSS. Backend is developed with python and Flask framework[6] with necessary libraries.

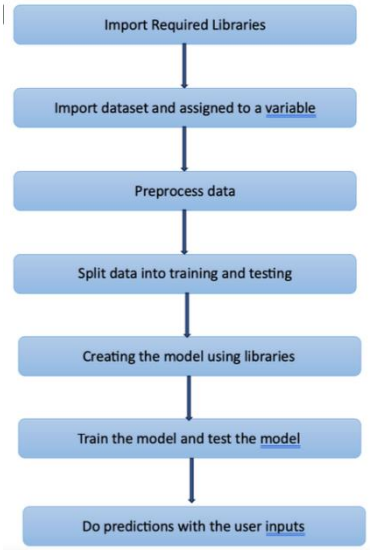


Figure 1. Designing predictive Algorithm.

At the beginning for the machine learning model, Naïve Bayes classification based model and Random Forest classification-based model [8] was built and compared the accuracy of two model using python programming language[9]. For the system implementation Naïve Bayes classification-based model [7] was used since it had the highest accuracy (93.58%) from both models. Designing and implementing both models were done in a same flow to compare accurately [14].

To visualize higher education sector details, google charts were used to create charts and graphs according to the selected details. Then it was embedded to the performance predicting web platform to be viewed.

IV. Evaluation And Results

predictive models were implemented using two statistical and machine learning approaches using Naïve Bayes classifier and Random Forest classifier [7] and then thoughtfully evaluated their effectiveness to choose most appropriate model for the building system.

When checking the accuracy and Mean Squared Error of the built two models, Naïve

Bayes classifier based model had 93.58% accuracy and Mean Squared Error was 0.103. When calculating the Mean Squared Error in model with Random Forest classification, it was 0.188. When considering the built two models, Naïve Bayes classifier used model has given the highest accuracy and lower mean squared error value after training and testing the model with fed dataset. Therefore, Naïve Bayes classifier used model was selected to implement the student performance prediction artifact.

Then the Naïve Bayes classifier-based models was evaluated using confusion matrix. Here the using the confusion matrix, the model was evaluated calculating the precision which interpret how many correctly predicted values are actually correct values. It has given the following accuracy levels for each subject which are in acceptable ranges.

Confusion matrix results	**Three main subjects**		
	Combined Mathematics	***Physics***	***Chemistry***
Precision	75%	92.3%	78.5%

Table 5. Confusion matrix results.

The main objective of this study is successfully accomplished at the last which was to build a machine learning based performance prediction model that predict students' A/L result using their previous three term test results with a good accuracy. It is anticipated to reduce the failure rate of physical science stream by the time with the developed predictive model. It is not going to be a direct outcome. But with the students start to use the system and get to know about their performance level then find the solution to improve their skills will gradually increase the passing rate and improve the overall performance of physical science stream.

The platform that visualizes about current university undergraduate counts, graduate counts, students count enrolled with higher education in government and non-government institutes were represented. Degrees and carrier paths available with them were not represented as supposed in the solution since it was a vast area that requires further studying.

V. Future Recommendations

As further recommendations, data visualization of higher education sector can be expanded by adding more relevant details according to the new requirements that may come up with the feedback of stakeholders or adding more areas that need to be disclosed such as available uncommon degree paths in state universities, career opportunities and graduate counts of each field that may helpful for the industry to recruit suitable employees.

VI. Conclusion

In summary, the proposed solution was implemented successfully to address designing a students' performance prediction model using machine learning to reduce high failure rate of physical science stream at GCE A/L examination as an online platform that will be easily predict performance and display for the user to have an understanding on their performance level and take other benefits of the platform provides. It is anticipated that the implemented solution will be a great impact on motivating students to prepare well for the examination and gradually the failure rate of the physical science stream will be decreased.

Acknowledgment

I would like to express my sincere gratitude to my supervisor, Mr. Gayan Perera for the continuous support for my final year research project for motivation, and guidance. Also, my gratitude goes to my family, my parents for their encouragement and support given to succeed in my academics.

References

[1] Lalith Edirisinghe, P. (n.d.). Leading the Children in Sri Lanka to Right Choice of Educa.on. h^ps://ssrn.com/abstract=3853658

[2] Kaushalya Abeywardhane, K. G., & Gamage, A. (2022). Subject Stream PredicWon: A Machine learning Approach to Select the Suitable Subject Stream for Senior Secondary Students in Sri Lanka. In Interna.onal Journal of Innova.ve Science and Research Technology (Vol. 7). www.ijisrt.com

[3] Dien, T. T., Luu, S. H., Thanh-Hai, N., & Thai-Nghe, N. (2020). Deep learning with data transformaWon and factor analysis for student

performance predicWon. Interna.onal Journal of Advanced Computer Science and Applica.ons, 11(8), 711–721. h^ps://doi.org/10.14569/IJACSA.2020.0110886

[4] Chauhan, N., Shah, K., Karn, D., & Dalal, J. (n.d.). Predic.on of Student's Performance Using Machine Learning. h^p://ssrn.com/link/2019-ICAST.html

[5] Yohannes Obsie, E., & Ahmed Adem, S. (2018). PredicWon of Student Academic Performance using Neural Network, Linear Regression and Support Vector Regression: A Case Study. In Interna.onal Journal of Computer Applica.ons (Vol. 180, Issue 40).

[6] Zainudin, A., Yunant, A. A., Politeknik Elektronika Negeri Surabaya, InsWtute of Electrical and Electronics Engineers. Indonesia SecWon, InsWtute of Electrical and Electronics Engineers, InternaWonal Electronics Symposium on Engineering Technology and ApplicaWons (2019 : Surabaya, I., & InternaWonal Electronics Symposium on Knowledge CreaWon and Intelligent CompuWng (2019 : Surabaya, I. (n.d.). Proceedings, IES 2019 : IES, Interna.onal Electronics Symposium : Surabaya, Indonesia, September 27-28, 2019 : the Role of Techno-intelligence in Crea.ng an Open Energy System Towards Energy Democracy.

[7] Sawant, N., & Khadapkar, D. R. (2022). Comparison of the performance of GaussianNB Algorithm, the K Neighbors Classifier Algorithm, the Logistic Regression Algorithm, the Linear Discriminant Analysis Algorithm, and the Decision Tree Classifier Algorithm on same dataset. International Journal for Research in Applied Science and Engineering Technology, 10(12), 1654–1665. https://doi.org/10.22214/ijraset.2022.48311

[8] Tripathi, A., Yadav, S., & Rajan, R. (2019). Naive Bayes Classification Model for the Student Performance Prediction. 2019 2nd International Conference on Intelligent Computing, Instrumentation and Control Technologies, ICICICT 2019, 1548–1553. https://doi.org/10.1109/ICICICT46008.2019.8993237

[9] Saabith, A. S., Vinothraj, T., & Fareez, M. (n.d.). A Review on Python Libraries and IDEs for Data Science. In International Journal of Research in Engineering and Science (IJRES) ISSN. www.ijres.org

[10] Madhawa, W. V. D. S., Supervisor, W., & Skrzypiec, G. (2019). Predicting Performance in the Sri Lankan General Certificate of Education Advanced Level Examination By Cognitive Abilities of the Test Takers Student.

[11] Department of Examinations, Sri Lanka. (n.d.).

[12] Sekeroglu, B., Abiyev, R., Ilhan, A., Arslan, M., & Idoko, J. B. (2021). Systematic literature review on machine learning and student performance prediction: Critical gaps and possible remedies. In Applied Sciences (Switzerland) (Vol. 11, Issue 22). MDPI. https://doi.org/10.3390/app112210907

[13] Institute of Electrical and Electronics Engineers., Pacific Distance Multimedia Education Network., & Mahāwitthayālai Sayām. (2012). Proceedings, 2012 Tenth International Conference on ICT and Knowledge Engineering : November 21-23, 2012 : Bangkok, Thailand. IEEE.

[14] Deo, R. C., Yaseen, Z. M., Al-Ansari, N., Nguyen-Huy, T., Langlands, T. A. M. P., & Galligan, L. (2020). Modern Artificial Intelligence Model Development for Undergraduate Student Performance Prediction: An Investigation on Engineering Mathematics Courses. IEEE Access, 8, 136697–136724. https://doi.org/10.1109/ACCESS.2020.3010938

[15] Albreiki, B., Zaki, N., & Alashwal, H. (2021). A systematic literature review of student' performance prediction using machine learning techniques. Education Sciences, 11(9). https://doi.org/10.3390/educsci11090552

Assessing the Effectiveness of a Digital Platform in Improving Citizen Engagement and Satisfaction with Public Transportation Services in Sri Lanka

Kreshan Jayamanna
Faculty of Computing
NSBM Green University
Pitipana, Homagama, Sri Lanka
kreshanjayamanna@icloud.com

Dulanjali Wijesekara
Faculty of Computing
NSBM Green University
Pitipana, Homagama, Sri Lanka
dulanjali.w@nsbm.ac.lk

Abstract*— In Sri Lanka, the prevalent preference for private vehicle transportation over public options. This preference results in critical problems such as traffic congestion, environmental pollution, and reduced fuel efficiency, impacting both the environment and the economy. Existing public transportation services suffer from quality issues, inconvenient routes, and unreliable schedules, contributing to citizen dissatisfaction. Moreover, The inadequate adoption of digital technology in Sri Lanka has further compounded these challenges, underscoring the pressing need for innovative solutions, This research undertook a comprehensive investigation into the root causes of the citizens' strong preference for private vehicles and their explicit expectations from public transportation systems in response to this, the study is conducted both quantitative and qualitative research to identify the reasons behind citizens' preference for private vehicles and their expectations from public transportation. The research involved feedback from citizens and transportation authorities. To understand this, we talked to around 140 people from different areas of Sri Lanka to figure out what they really want from public transport. The findings underscore the need to enhance citizen engagement with public transportation through the development of a tailored system. This system comprises a mobile app for passengers, bus drivers, conductors, and owners, as well as a web application for administrative and transportation board authorities. It offers accurate information on public transportation locations, fares, and feedback mechanisms, aligning with citizens' expectations. Although this system currently focuses on a limited number of routes in Sri Lanka, it has the potential for further expansion. By obtaining more input from the people of Sri Lanka and implementing their suggestions, the application can be improved and adapted to serve all routes in the country. In essence, this research aims to alleviate the problems associated with private vehicle transportation and enhance public transportation's appeal, offering a practical solution for the transportation challenges faced in Sri Lanka.

Keywords*—Citizen engagement, Digital technology adoption, Information accuracy, Feedback mechanisms, Citizen preferences

I. INTRODUCTION

Transportation is an essential service that arises from the need to move people and goods from one location to another. It involves the transfer of passengers and various types of goods between a starting point and a desired destination, utilizing different methods of conveyance, public transportation is a shared passenger transportation service that is available to the general public for use in getting from one place to another [1], Public transportation plays a crucial role in providing convenient and sustainable mobility options for citizens in Sri Lanka. However, there is a growing concern regarding citizens' preference for private vehicle transportation over public transportation. This research aims to evaluate the effectiveness of a digital platform in improving citizen engagement and satisfaction with public transportation services in Sri Lanka. By exploring the

current state of public transportation, identifying reasons for citizens' preference for private vehicle transport, and exploring opportunities for improvement in engagement and satisfaction with public transport, this study seeks to address the challenges and potential solutions to enhance the public transportation experience.

In recent years, there has been a significant increase in private vehicle ownership in Sri Lanka, leading to issues such as traffic congestion, environmental pollution, and a strain on existing infrastructure. Understanding the factors contributing to citizens' preference for private vehicle transport is crucial to identify the barriers to public transportation 2 usage and formulate effective strategies for improvement. The advent of digital platforms has the potential to transform the way citizens interact with public transportation services. By leveraging technology, it is possible to enhance engagement, convenience, and overall satisfaction with public transportation. The research will start by examining the current state of public transportation in Sri Lanka, considering factors such as the bus services' coverage, affordability, accessibility, and reliability. It will also explore the existing challenges faced by users, such as overcrowding, irregular schedules, and limited connectivity, This study provides a number of issues with bus transportation in Sri Lanka, including significant and rising deficits, poor operational practices, and poor management [2]. These insights will provide a baseline understanding of the public transportation system's strengths and weaknesses. To identify the reasons for citizens' preference for private vehicle transport, the research will delve into factors such as convenience, comfort, flexibility, speed, and safety concerns. By understanding these underlying motivations, it will be possible to address the gaps and design interventions that align public transportation services more closely with citizens' needs and preferences.

Furthermore, the study will assess the opportunities for improving engagement and satisfaction with public transport through the implementation of a digital platform. It will examine the potential benefits of real-time information dissemination, personalized trip planning features, and mechanisms for citizen feedback and involvement. The research will also consider the challenges and limitations of implementing a digital platform in a diverse and evolving transportation landscape. By conducting this research, my is aim to provide valuable insights into the current state of public transportation in Sri Lanka, the reasons behind citizens' preference for private vehicle transport, and the potential for improving engagement and satisfaction with public transport through the adoption of a digital platform. The findings will inform policymakers, transportation authorities, and stakeholders in developing strategies and initiatives to enhance the public transportation experience and promote sustainable mobility options in Sri Lanka.

A. Research Problem

The preference for private vehicle transport in Sri Lanka hinders sustainable urban mobility and negatively impacts public transportation. Efforts to improve public transport services have not effectively engaged citizens or increased satisfaction. A lack of research on digital platforms to enhance citizen engagement and identify desired features and varying impacts across citizen groups exists. Addressing these gaps is vital for improving public transportation and promoting sustainable mobility in Sri Lanka.

B. Research Question

How can a digital platform be designed, optimized, and implemented to effectively enhance citizen engagement and satisfaction with public transportation services in Sri Lanka, considering the factors that contribute to citizens' preference for private vehicle transport over public transportation, the key features desired by citizens, and potential barriers to adoption and use of the platform?

C. Research Objectives

- To identify the reasons caused citizens in Sri Lanka prefer private vehicle transport over public transportation services.
- To assess the level of citizen engagement and satisfaction with public transportation services in Sri Lanka.
- To identify potential barriers to adoption and use of the digital platform and to develop strategies to overcome these barriers and to explore the key features desired by citizens for a digital platform that can improve their engagement and satisfaction with public transportation services in Sri Lanka.
- To design and implement a digital platform that can effectively enhance citizen engagement and satisfaction with public transportation services in Sri Lanka.
- To evaluate the effectiveness of the digital platform in increasing citizen engagement and satisfaction with public transportation services in Sri Lanka.

II. Literature Review

A. Importance of public transportation in Sri Lanka

In Sri Lanka, public transportation is extremely important and provides a variety of advantages that greatly improve the economic growth and general well-being of the nation. One of its main benefits is the ease of access and low cost it provides to the citizens ensuring that people from all walks of life can access education, healthcare, employment, and other essential services. [3] have mentioned in the study that he has conducted that public transportation is a crucial component of Sri Lanka's infrastructure and offers its citizens a number of advantages. The primary form of public transportation is the bus, which is operated by both privately owned buses and the government-run Sri Lanka Transport Board (SLTB). This inclusiveness is especially important in a nation where a significant percentage of the population might not have enough money for buying their own vehicles. For millions of people, public transportation turns into a lifeline that enables them to actively engage in society and the economy. Public transportation in Sri Lanka is crucial in that it has the ability to address the critical issue of traffic congestion and its adverse effects on the environment. Cities in Sri Lanka face significant congestion, which increases travel times, consumes more fuel, and produces harmful pollution. One of the significant issues in urban areas, especially in major cities like Colombo, is traffic congestion, by encouraging more people to take public transportation and lowering the number of private vehicles on the roadways, an effective and well-designed public transportation system helps ease this congestion. As a result, there are less carbon emissions and air pollution, which helps to create a more livable and sustainable environment.

Public transportation provides significant economic advantages in addition to its environmental advantages. An efficient system improves productivity by providing seamless mobility options for the workforce, reducing commuting time, and avoiding the stress of navigating traffic congestion. This greater mobility helps in improving overall productivity and strengthening the economy [4], Additionally, enhanced interaction between urban and rural areas is encouraged through public transportation, resulting in regional development and economic expansion. An efficient public transportation system can have a big impact on how cities develop in terms of urban planning and land use. An urban planning idea known as

transit-oriented developments (TODs) maintains a focus on small, accessible regions centered around public transportation hubs as well as mixed-use development structures. Adopting TODs can enable more effective land use, limit urban sprawl, and promote sustainable urban growth [4]

The development and well-being of Sri Lanka is greatly affected by public transportation. Additionally, to improving citizens accessibility and mobility, it promotes social equality, environmental sustainability, and economic prosperity. Sri Lanka may develop a more productive, inclusive, and resilient society for the benefit of all its citizens by giving public transportation infrastructure and services priority and investment.

B. Importance of using technology in transportation

The integration of technology in transportation revolutionizes operations, promoting efficiency, customer satisfaction, and seamless connectivity between transportation modes. This includes digital networks, mobility platforms, real-time analytics, and intelligent systems, all aimed at optimizing routes and enhancing operational effectiveness [5].

Technological advancements in transportation lead to cost reduction and operational efficiency through automation and digitization. Predictive maintenance, facilitated by analytics and machine learning, prevents costly breakdowns. Real-time data aids inventory management, logistics, and faster deliveries, while GPS tracking and telematics enhance passenger satisfaction and service predictability.

Moreover, technology enables coordinated and multifunctional travel options by connecting various transportation modes. It empowers proactive decision-making based on real-time data, promoting safety, sustainability, and effective infrastructure planning. The culmination of these advancements results in a more efficient and evolved transportation industry.

C. Challenges faced by citizens using public bus transportation in Sri Lanka

Sri Lanka's public transportation confronts numerous challenges affecting citizens' daily lives [6]. Inconsistent schedules, overcrowding, poorly maintained vehicles, and law-breaking drivers create inconvenience and safety concerns for passengers [2]. Furthermore, traffic congestion exacerbates delays and inefficiencies in both public and private transportation systems [7].

In addition to these challenges, inadequate facilities and cleanliness in stations and vehicles add to the discomfort of commuters, particularly in urban areas [6] ,Insufficient seating, cracked windows, and the lack of air conditioning make travel unpleasant, especially in adverse weather conditions. The lack of basic amenities impacts passenger experience and satisfaction.

To address these multifaceted challenges, Sri Lanka necessitates a comprehensive strategy: increased funding for infrastructure, improved maintenance standards, enhanced service frequency, strict policies ensuring safety and accessibility, and investment in modernized facilities [7]. Addressing these issues holistically would significantly uplift the quality and efficiency of Sri Lanka's public transportation, providing its citizens with

a more reliable and comfortable commuting option.

D. *Public and Private vehicle usage in Sri Lanka*

In Sri Lanka, buses, both state-owned and private, form the primary public transportation system, offering various comfort, cost, and speed options [3] However, the system faces challenges, including unequal bus availability across routes, resulting in delays and imbalances in income for owners. This, coupled with rising private vehicle ownership due to expectations of efficiency and convenience, leads to traffic congestion and environmental issues [3], While statistics show a higher usage of private vehicles, a decline in public transportation usage has been observed over the years, mainly due to dissatisfaction with service quality and inefficiencies [1]. As passenger satisfaction declines, there's a shift towards personal vehicles, reflecting a decrease in public transportation use over time. To encourage public transportation use, improving service quality, meeting diverse customer needs, and adhering to bus standards are crucial factors [2].

E. *Factors influencing mode choice in transportation*

Transportation mode selection in Sri Lanka involves a choice between private vehicle ownership and utilizing the public bus system, influenced by various factors. Understanding consumer preferences aids decision-makers in enhancing transportation services to align with customer demands. Comfort, reliability, cost-effectiveness, safety, convenience, and environmental sustainability are key attributes to consider in creating a transportation system that satisfies consumer expectations [8]. Sri Lanka's population of 20.8 million is spread across nine provinces, with the Western Province, housing Colombo, the commercial hub. The country's vehicle ownership rate is 129 vehicles per 1000 people, but for cars, vans, and three-wheelers, it's around 46 vehicles per 1000 individuals [9]. This demonstrates a significant presence of motor vehicles in the country, highlighting the need for robust transportation infrastructure and services.

The choice between public bus transportation and private vehicles in Sri Lanka hinges on various factors. Private vehicles offer flexibility and convenience in scheduling and routes, while buses provide broad accessibility. Cost is a crucial consideration, with private ownership being more expensive compared to cost-effective bus fares. Time efficiency varies between the modes, with private vehicles offering faster travel but buses potentially faster in congested urban areas. Environmental impact favors buses due to their sustainability. Accessibility is vital for those without private vehicles, and safety considerations differ between buses and individual vehicle responsibility. Road infrastructure, personal preferences, traffic congestion, and parking availability also shape transportation choices [2]. Policymakers should focus on enhancing public bus quality, efficiency, and reliability to encourage a shift towards more sustainable transportation modes while considering diverse population needs.

F. *Citizen Engagement and Satisfaction Improvement related Frameworks and related theories and concepts*

- Technology Acceptance Model (TAM): TAM evaluates how people perceive and use new technology, focusing on usability and utility [10] . It's valuable for understanding digital platform adoption in public transportation and guiding policy decisions. [11] note its simplicity and applicability

in explaining technology acceptance.

- Citizen-Centric Service Delivery Model: This model prioritizes citizen engagement and satisfaction in public [12]. Assessing the platform's adherence to citizen-centric principles helps gauge its effectiveness in promoting citizen involvement. Strategies should focus on understanding public needs for service transformation [13]emphasize the importance of citizen involvement in successful e-government services.

G. Existing Systems

Design and Development of Android Mobile Based Bus Tracking System [14] The Mobile-Based Bus Tracking System is an Android app that efficiently tracks buses in real-time using GPS and GSM technologies. Passengers and coordinators can access the bus location, estimated arrival time, and distance from their stop via internet-enabled Android phones. This solution minimizes communication interruptions, enhances convenience, and offers reliable and adaptable bus tracking.

Public Bus Tracking System for Sri Lanka [6] This mobile and web-based system tracks public transportation to address issues like congestion, delays, and unreliable schedules. It utilizes GPS for bus schedules, real-time tracking, and travel planning. Admins manage the system, update bus info, and handle issues. Passengers access the app for bus details and arrival estimates, while drivers track their buses through GPS.

Real Time Web-Based Bus Tracking System [15] The mobile application developed as part of this system offers a comprehensive solution for users to access relevant information about buses operating between their specified source and destination. In order to detect the buses' current location and provide position coordinates with latitude and longitude, the application employs GPS equipment mounted on the buses. Utilizing wireless communication networks like GSM/GPRS, the tracking data is transferred. Remote users can visit the mobile application to get the most recent details on bus numbers, route information, and the actual location of the buses that are going between their designated source and destination.

H. Gaps Identified in existing systems

To fully comprehend why Sri Lankans, prefer private vehicles to public transportation, especially in regard to digital technology, a thorough research study is required. This study should look into aspects of both forms of transportation that relate to comfort, reliability, punctuality, privacy, and cost considerations. To support sustainable and effective transportation systems in Sri Lanka, plans and actions can be informed by a deeper awareness of citizens' preferences and the potential role of digital technology in public transportation.

The absence of support for the Sinhala and Tamil language in different established systems, which predominantly prioritize the English language, has been noted as a gap. To close the language gap and advance inclusivity, it is essential to integrate support for the Sinhala language in digital systems and technology. For a sizeable segment of the population, this would enhance usability, accessibility, and user experience, enabling efficient participation and engagement with digital platforms.

Limited Geographic Representation in Data Collection for Sri Lanka and Lack of Feedback Mechanisms and Lost Belongings Assistance functionalities in existing systems in Sri Lanka.

III. Research Methodology

A. Research Paradigm

This research study focused on enhancing citizen engagement and satisfaction with public transportation systems in Sri Lanka, research will be adopting a mixed paradigm approach that combines both positivism and interpretivism. This comprehensive approach enables to capture both quantifiable patterns and qualitative insights, ensuring a well-rounded understanding of the complex dynamics surrounding the problems.

B. Research Apporach

This research will be following deductive research approach, the deductive research approach starts with well-established theories, concepts, or models or the theories which already exits and then moves towards empirical observations and data collection to test and validate these theories. In particular, the Citizen-Centric Service Delivery Model will be utilizing the as a theoretical framework to guide this research. This framework provides a structured framework for understanding and assessing the factors that contribute to citizen satisfaction and engagement in the realm of public transportation.

C. Research Stratergy

The action research approach is adopted as the research strategy, the action research methodology attempts to both analyze and resolve contemporary problems. This strategy, as its name suggests, entails conducting research while actively taking significant steps to address the problems discovered. In the context of this research, which focuses on enhancing citizen engagement and satisfaction in Sri Lanka's public transportation system, the Action Research approach highly relevant.

D. Data collection mechanisms

This research aims to gather extensive and diverse data on Sri Lanka's public transportation system by employing various data collection techniques. These methods are carefully chosen to encompass the perspectives of different stakeholders, including passengers, citizens, officials, and industry workers. Through questionnaires and interviews, the research seeks to create a comprehensive understanding of current challenges, opportunities, factors influencing citizen satisfaction, and areas for improvement in the public transportation system. By using both quantitative and qualitative approaches, the research aims to triangulate the data, ensuring accuracy and reliability in the research findings.

The research will utilize questionnaires targeting passengers and citizens using Sri Lanka's public transportation. This approach aims to create a representative sample encompassing diverse backgrounds, ages, and locations, ensuring inclusivity. The questionnaires, available in Sinhala and English via Google Forms, aim to engage a broad participant base. Employing SurveyMonkey helped determine an appropriate sample size of over 96 responders, balancing reliability and practicality in data collection.

Interviews is used to engage with key players within Sri Lanka's public transportation sector, focusing on officials from the Sri Lanka Central Transport Board (SLTB), drivers, and private bus owners. Qualitative insights will be obtained through in-depth interviews, intending to delve into the operational intricacies, challenges, viewpoints, and experiences of these stakeholders.

E. Data Sources

The research focuses on Sri Lanka's public transportation system using both

primary and secondary sources. It utilizes questionnaires and interviews to gather data from passengers and key stakeholders like SLTB officials and bus operators. This multifaceted approach aims to understand user experiences, identify trends, and develop strategies to enhance citizen engagement and satisfaction with the transportation system.

IV. System Requirement Specification

A. *Research Onion Model*

The research onion by Saunders is a widely used framework in social sciences and business research, depicting phases and levels in the research process. Methodology serves as a comprehensive research approach guiding procedures, principles, and philosophical perspectives [16]. It's a fundamental aspect in dissertations and theses, ensuring alignment between tools, techniques, and the underlying philosophical framework.

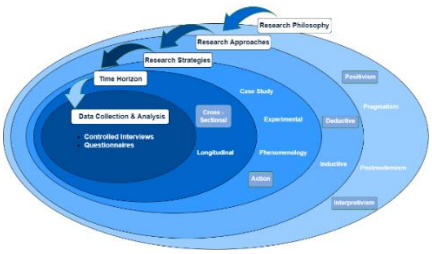


Figure 1. Research Onion

Stakeholders	Data Collection Method
General Public	Questionnaires
SLTB Officials	Controlled Interviews
SLTB Drivers/Conductors	Controlled Interviews
Private Bus Owners	Controlled Interviews

Table 1. Stakeholders and how data are collected from them

B. *Interviews Analysis Under Thematic Analysis*

Thematic analysis is a qualitative research method that is widely used in the social sciences, psychology, and other fields. Its major goal is to explore and identify patterns, themes, and meanings within a given dataset, which often include textual or visual information[17] This method provides investigators with a versatile and systematic way to delve into the substance of their data, allowing them to uncover underlying themes and patterns while gaining a more profound understanding of the content.

C. *Questionnaires Analysis Under Descriptive Statistics*

The data analysis conducted based on the questionnaires that were administered in both Sinhala and English versions. These questionnaires contained identical questions in both languages. To perform the analysis, here the responses have been combined the total responses from both the Sinhala and English versions and calculated various percentages and other relevant statistics. This combined dataset comprises a total of 146 responses from participants who completed the questionnaires in either language.

D. *Interview Results*

The analysis of gathered insights revealed several key findings regarding transportation preferences and challenges faced by citizens. Primarily, the appeal of private vehicles remains high due to their convenience, flexibility, and perceived reliability, factors that strongly influence people's choice of transportation. However, a widespread issue reported by interviewees was the overcrowding experienced during peak hours on public buses, causing discomfort and inconvenience for passengers across the board.

Interestingly, both the state-owned SLTB (Sri Lanka Transport Board) and private bus owners expressed interest in leveraging

digital applications to address existing issues and enhance overall service quality. There was a shared optimism among interviewees regarding the potential for improvement through digital initiatives. They believed that advanced technological solutions could effectively engage citizens and bring about substantial enhancements in the transportation system. This collective sentiment points towards a willingness to embrace technology as a means to create a more efficient and engaging transportation experience for all.

E. Questionaries' Results

The analysis of gathered data revealed insightful patterns regarding public transportation usage and preferences among respondents. A notable 33% of participants acknowledged frequent use of public transportation, indicating a significant portion relying on these services. A substantial majority, accounting for 82% of respondents, utilize both buses and trains, with the remainder opting for alternative modes of transport.

Key factors influencing transportation choices emerged prominently among most participants. Time efficiency, availability, and cost were cited as primary considerations when selecting a mode of transportation. However, challenges in the public transportation system were also highlighted, including issues like unreliable schedules, insufficient fare information, limited routes, delays, and overcrowding. These factors collectively contribute to an inconvenient experience for citizens relying on these services.

Regarding improvements desired in public transportation, respondents expressed a pressing need for enhanced facilities, reduced time delays, better bus quality, improved time management, and an overall system upgrade within Sri Lanka's transportation infrastructure. Notably, an overwhelming 94% of respondents expressed a keen interest in using public transportation if provided with user-friendly digital platforms offering real-time information and other valuable features. This strong endorsement underscores the potential impact of technology in significantly increasing public transportation usage among citizens if appropriately implemented.

F. Proposed system architecture diagram of application Design

Figure 2. Proposed System

The document provides an overview of a proposed system architecture for an application designed to enhance citizen engagement and satisfaction with public transportation services in Sri Lanka. It includes a diagram (Figure 2) illustrating the system's design and mentions the use of various technologies for backend and frontend development, such as the Spring Boot Framework for backend, Angular Framework for the NaviGo Admin Portal Web Application, and Ionic Framework & Angular Framework for the NaviGo Mobile Application. Additionally, it details the use of PostgreSQL for the database, AWS RDS for database hosting, AWS EC2 for backend

deployment, and AWS S3 for web application deployment.

V. Design & Implementation

A. Backend & Frontend Technology Usage

Technology	Description
Spring Boot Framework	Backend Development
Angular Framework	Frontend NaviGo Admin Portal Web Application
Ionic Framework & Angular Framework	Frontend NaviGo Mobile Application

Table 2. Backend & Frontend Technology used for development

B. Database & Hosting Technology Usage

Technology	Description
PostgresSQL	Relational Database
AWS RDS	Database is running under AS RDS PostgreSQL
AWS EC2 Instance	Backend Deployment
AWS S3 Bucket	Angular Web Application Deployment

Table 3. Database & Hosting Technologies usage

C. NaviGo Mobile Application UI

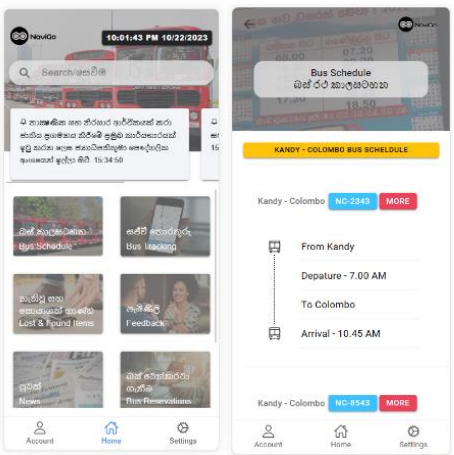


Figure 3. NaviGo Mobile Application

The home screen figure 03. provides a clean and intuitive interface, featuring a navigation menu for quick access to all key functionalities such as searching for bus routes, tracking buses in real-time, and viewing bus schedules. It might also include a user profile icon for easy access to account settings and preferences, and possibly shortcuts to frequently used services like favorite routes or ticket purchases.

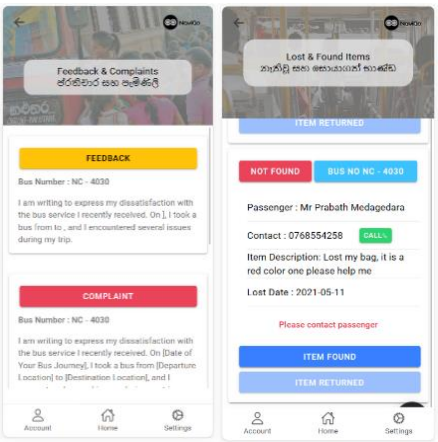


Figure 4. NaviGo Mobile Application

This screen figure 04. probably allows users to input their starting location and destination to find the best available routes. It might display options including the shortest route, the least number of transfers, or the fastest journey based on current traffic and bus schedules. Each route option could provide details like estimated travel time, bus numbers involved, and the number of stops or transfers.

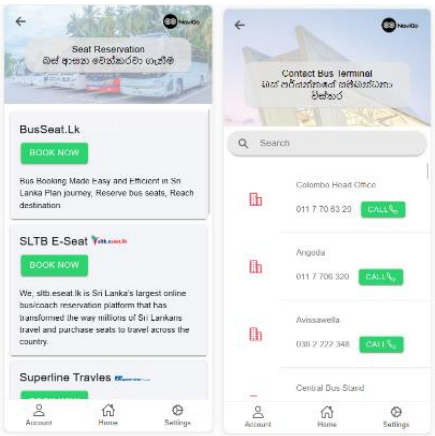


Figure 5. NaviGo Mobile Application

Here in figure 05, users can likely view the real-time locations of buses on a map, including details such as the estimated time of arrival (ETA) at their chosen stop. This feature might help users plan their departure to minimize waiting times and stay updated on any delays or changes in the schedule.

D. NaviGo Admin Portal Web Application UI

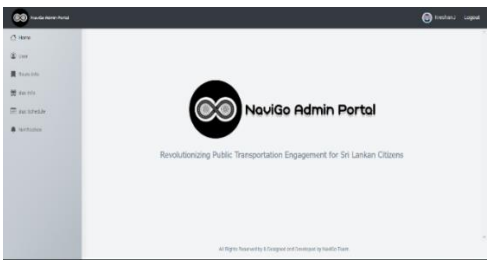


Figure 6. NaviGo Admin Portal

This figure06 part of the application might allow users to manage their personal information, preferences, and any saved data such as favorite routes or stops. Settings for notifications, language, and accessibility options could also be included to tailor the app experience to individual needs. This screen figure 07 could encompass various supplementary features such as a fare calculator, customer service or feedback forms, and information on bus capacities or amenities (e.g., Wi-Fi, air conditioning). It might also offer community-driven content like ratings or reviews of bus routes and services.

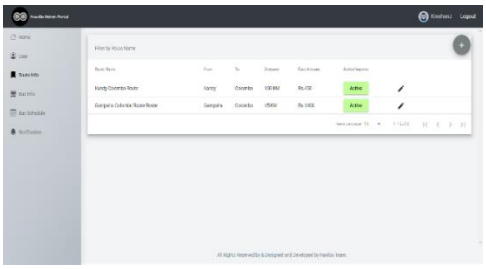


Figure 7. NaviGo Admin Portal

VI. CONCLUTION

This research main objective was to understand why citizens in Sri Lanka prefer private vehicles over public transportation. It identified convenience, accessibility, cost, and comfort as key factors driving this preference. Addressing these issues can make public transport more attractive, reducing reliance on private vehicles and improving overall sustainability.

Surveys assessed citizen satisfaction with existing public transport services, pinpointing areas needing improvement. This sets a baseline for future interventions to enhance public transportation. Feedback from citizens guided the development of a user-centric digital platform. It focuses on real-time info, trip planning, live bus tracking, feedback mechanisms, and lost item assistance, aligning closely with citizens' needs and enhancing their transportation experience.

The ongoing evaluation of the digital platform's effectiveness involves continuous user feedback and usage analysis. Positive user responses indicate the platform's relevance in addressing transportation concerns, shaping a more efficient and passenger-oriented public transportation

system in Sri Lanka. The research is centered around enhancing public transportation experiences in Sri Lanka. The fundamental idea behind this research is to address the issues people face with the current transportation services and create a digital platform that caters to their needs. Improving public transportation will have a positive impact on people's lives. My approach relies on user-centric approach. This research was intended to identify the problems with the transportation system and come up with practical solutions. The ultimate objective is to improve the convenience and satisfaction of using public transportation. The primary objective is to create a user-friendly digital platform that connects passengers and transportation services. This platform will encourage citizen engagement and satisfaction, which will ultimately result in a more effective and accessible transportation system for Sri Lanka.

A. Future Recommendations

The incorporation of multi-modal mobility can be done to increase the usefulness of the digital platform. While the current design focuses on bus transportation, adding rail schedules, routes, and real-time data would provide citizens with a complete, all-in-one solution.

Enable users to contribute real-time information about traffic conditions, delays, accidents, and other relevant data. This crowd-sourced information can help fellow travellers make informed decisions and improve overall travel experiences.

Implement a seamless and contactless payment system within the app, allowing users to pay for their transportation tickets or fares digitally. This will reduce the reliance on physical tickets and streamline the payment process.

REFERENCES

[1] H. Ranawana and D. Hewage, "Factors Affecting Service Quality in Public Bus Transportation in Sri Lanka," 2015.

[2] A. S. Kumarage, "Developing Public Transport in Sri Lanka Comparison of Speed data by Different Speed Detection Techniques View project Transport Demand Estimation with Cell Phone Data View project," 2005, doi: 10.13140/RG.2.2.33423.79528
.

[3] C. T. Danthanarayana, "Evaluation of Public Transportation Mode Choice with Passenger Satisfaction: A Case Study of the Semi-Luxury Bus Service on the Panadura-KandyRoute in Sri Lanka," 2019.

[4] [4] J. Hine, "Poverty and sustainable transport How transport affects poor people with policy implications for poverty reduction A literature review Paul Starkey Consultant in integrated transport," 2014.

[5] [5] T. Kanagasabai, N. Nishanthan, K. Thiruthanigesan, and P. Georgakis, "Towards The Development Of Intelligent Transportation Systems In Sri Lanka," Article in International Journal of Scientific & Technology Research, vol. 6, 2017.

[6] [6] W. Wijesinghe, D. Gunasekera, and M. RMM Pradeep, "Public Bus Tracking System for Sri Lanka."

[7] [7] S. Lebbe, A. Haleem, and S. Sabraz Nawaz, "Real Time Bus Tracking and Scheduling System Using Wireless Sensor and Mobile Technology," Journal of Information Systems & Information Technology (JISIT), vol. 1, no. 1, 2016.

[8] V. A. Arroyo et al., "TRANSPORTATION RESEARCH BOARD 2015 EXECUTIVE COMMITTEE*."

[9] R. A. M. MADHUWANTHI, A. MARASINGHE, R. P. C. J. RAJAPAKSE, A. D. DHARMAWANSA, and S. NOMURA, "Factors Influencing to Travel Behavior on Transport Mode Choice," International Journal of Affective Engineering, vol. 15, no. 2, pp. 63–72, 2016, doi: 10.5057/ijae.ijae-d-15-00044.

[10] H. Aulawi, "Technology Acceptance Model for Online Transportation," International

Journal of Advanced Trends in Computer Science and Engineering, vol. 9, no. 1, pp. 31–35, Feb. 2020, doi: 10.30534/ijatcse/2020/06912020.

[11] E. B. Diop, S. Zhao, and T. Van Duy, "An extension of the technology acceptance model for understanding travelers' adoption of variable message signs," PLoS One, vol. 14, no. 4, Apr. 2019, doi: 10.1371/journal.pone.0216007.

[12] E. Dudley, D.-Y. Lin, M. Mancini, and J. Ng, "Implementing a citizen-centric approach to delivering government services."

[13] Z. Wu, H. Zhou, H. Xi, and N. Wu, "Analysing public acceptance of autonomous buses based on an extended TAM model," IET Intelligent Transport Systems, vol. 15, no. 10, pp. 1318–1330, Oct. 2021, doi: 10.1049/itr2.12100.

[14] K. Sujatha, P. V. N. Rao, K. J. Sruthi, and A. A. Rao, "Design and development of android mobile based bus tracking system," in 1st International Conference on Networks and Soft Computing, ICNSC 2014 - Proceedings, Institute of Electrical and Electronics Engineers Inc., Sep. 2014, pp. 231–235. doi: 10.1109/CNSC.2014.6906694.

[15] M. Kumbhar, M. Survase, P. Mastud, A. Salunke, and A. Professor Shrinivas Sirdeshpande, "Real Time Web Based Bus Tracking System," International Research Journal of Engineering and Technology, 2016.

Simplified and Accurate License Plate Character Recognition System

Kanishka Dharmasena
Department of Computer Science
NSBM Green University
Homagama, Sri Lanka
sahankanishkaofficial@outlook.com

Abstract*—The rise of AI is transforming numerous industries including law enforcement and surveillance. Many businesses and homeowners now use advanced surveillance cameras but challenges such as poor lighting, paying a large sum of money and privacy concerns in commercial applications puts a hold on to these applications. Access to a LPR program is a valuable tool, enabling accurate extraction of license plate data from images and videos has proved crucial. LPR technology has a wide range of applications ranging from personal to government level requirements. In areas where vehicle theft is a considerable threat such LPCR systems can assist in identifying stolen vehicles and assisting law enforcement. The program in discussion provides a free to use LPCR which is network independent and privacy conscious. The system is simplified, free to use, does not require an internet connection and delivers accurate results with a accuracy of 87.5% by subjecting the user submitted still images to optical character recognition.

Keywords*—License Plate Character Recognition. OCR, Computer Vision, Data Privacy

I. INTRODUCTION

With the rise of technology and increased automations all around the world, we can observe that the widespread use of artificial intelligence is on the rise. Already some major global institutions and multinational conglomerates are using AI to ease their work.

One of the fields AI is used but not a lot discussed is law enforcement and general surveillance. The reason AI powered surveillance is not very popular is privacy concerns of the public. In these modern times most businesses and a significant percentage of homeowners have installed camera systems for protection against crimes. These new generation surveillance cameras are getting more sophisticated and produce more quality images than ever before. However, sometimes us humans have a hard time identifying vehicle license plates because of some complications such as bad lighting, quality of the image and presence of occlusions in the still image taken from the surveillance video.

In these recent times, computer vision and image processing has become mainstream with the outbreak of social media and the dawn of the age of influencers. We can use them for recreational purposes such as photo editing and enhancing and also for more serious functions. If we can access a program that helps us identify license plates, it would be of great help. The use of computer vision techniques and machine learning algorithms has made it possible to recognize and extract license plate information from digital images or video streams with high accuracy.

License plate recognition (LPR) is a critical technology that we can use for various purposes. We can see that the Peoples Republic of China has already implemented widespread use of LPR systems integrated through their surveillance. If an LPR system is made so that its easily accessible, we can use it in events such as traffic monitoring, law enforcement, parking management and toll collection (Example: ETC toll payments

in expressways). Given the socio-economic state Sri Lanka is in we have experienced a rise of grand theft auto. We can use the LPR system to identify stolen vehicles through their license plates thus helping law enforcement agencies.

Image processing is a method to perform some operations on an image, in order to get an enhanced image or to extract some useful information from it. It is a type of signal processing in which input is an image and output may be image or characteristics/features associated with that image. Nowadays, image processing is among rapidly growing technologies. It forms core research area within engineering and computer science disciplines too.

Image processing basically includes the following three steps:

• Importing the image via image acquisition tools;

• Analyzing and manipulating the image;

• Output in which result can be altered image or report that is based on image analysis.

There are two types of methods used for image processing namely, analogue and digital image processing. Analogue image processing can be used for the hard copies like printouts and photographs. Image analysts use various fundamentals of interpretation while using these visual techniques. Digital image processing techniques help in manipulation of the digital images by using computers. The three general phases that all types of data have to undergo while using digital technique are pre-processing, enhancement, and display, information extraction

A. Study Area and Study Objectives

The **study area** of this project is the domain of computer vision and image processing specifically in the development of a license plate recognition program.

The **domain** of the proposed research project focuses on the domain of computer vision and image processing. This research also aims to explore the current state of the art techniques and algorithms for pattern recognition.

The **study objective** is to study image processing and machine learning methodologies and implement a robust and an efficient license plate recognition system that can accurately detect and recognize the license plates in either good or bad lighting and weather conditions. I also plan to investigate the performance of different machine learning algorithms and object identification models and compare them to identify what the optimal solution is when creating a light weight and easy to use license plate recognition system and compare the effectiveness against real life scenarios.

B. Technology

Using OpenCV libraries and Tesseract OCR methods.

In simple terms, OpenCV is a computer software library that helps computers see, understand, analyze images and videos. Its similar to a toolbox that contains various types of tools that allow developers to build computer vision relate programs. OpenCV is proven to be useful in many instances since its release. Another major fact about OpenCV is its open source (meaning its free to use and anyone can contribute to its development making the proposed program open for upgrades and refinements when the image processing technology evolves and proving it has the potential to be a powerful tool for computer vision applications)

OpenCV (Open Source Computer Vision Library: http://opencv.org) is an open-source library that includes several hundreds of computer vision algorithms. The document describes the so-called OpenCV 2.x API, which is essentially a C++ API, as opposed to the C-based OpenCV 1.x API (C API is deprecated and not tested with "C" compiler

since OpenCV 2.4 releases (Documentation, n.d.))

The proposed license plate recognition system aims to explore the correlation between computer vision techniques and machine learning algorithms. This research will investigate the effectiveness of different computer vision techniques such as edge detection, morphological operations and image segmentation in preprocessing the images given as an input and extracting the relevant license plate regions.
Image processing is a robust involving diverse methodologies. The system uses series of image processing techniques to detect, segment and recognize the vehicle license plate. However, it should be noted that a complete license plate recognition system requires effective set of hardware and software components, most preferably efficient infra-red cameras and powerful computers to provide high quality images. Canny edge detector as it is proven to have higher accuracy.

C. Research Gaps and Research Problems

Even though there are some license plate recognition systems, there are only a few available out to the public and there is even fewer LPR systems that provide their services free of charge. There are also issues with some when it comes to accuracy to due varying lighting and weather conditions. Research should be conducted on how to improve the accuracy of these systems in such conditions.

Another important fact is that there is less information on how to integrate the LPR system with other technologies such as integrating the program with surveillance cameras or GPS tracking. Research should be conducted on regards to how to integrate such system with a LPR system. According to the hypothesis we should be able to integrate the proposed system to external hardware using a Raspberry Pi. Using this method, we can integrate the LPR with a number of electronic surveillance hardware.

A reason why there is a lack of documentation regarding integrating a LPR with general surveillance is because of privacy related concerns of the public because the system can capture sensitive information about drivers and their vehicles which could raise concerns about privacy and data protection. Due to this reason law enforcement agencies and governments tend to make detailed documentations of such systems publicly available.

Another issue that concerns this project is whether the proposed model be able to identify a wide range of license plates because these plates differ from county to county and also region-wise. This factor may affect the accuracy of recognition. Research should be conducted in order to explore how to design license plate recognition systems that can accurately recognize plates from a range of regions.

The research analysis conducted by Joao Manuel R.S. Tavares states the challenges associated with analyzing objects in images. He has mentioned various tasks involved in the analysis such as segmentation, feature extraction, image matching, alignment, tracking, motion analysis. This paragraph below was quoted from one of his publishing which highlights the difficulty of carrying out these tasks in an automated and efficient manner but also notes that some of the tasks are interrelated. (MMP Petrou, C Petrou, 2010)

"The computational analysis of objects in images is a very challenging issue as it usually involves automatic tasks for segmentation, that is, the detection of the objects represented, extraction of representative features from the objects, matching between images, rigid and non-rigid alignment of images, temporal tracking

and motion analysis of features in image sequences, deformation estimation between two objects, as well as the 3D shape reconstruction of the objects from these images. Although, to carry out each of these tasks in a fully automatic, efficient and robust manner is generally demanding, some of these tasks often appear associated. For example, to analyze the behavior of organs from sequences of medical images, first the input images should be segmented, then suitable features of the organs under analysis should be extracted and tracked along the image sequences and finally the motions involved should be tracked and analyzed.

II. TECHNOLOGICAL ANALYSIS

A. Algorithmic Analysis

The algorithm for license plate detection aims to identify the presence and location of license plates within an image. Technologies like edge detection, template matching, and machine learning-based approaches are used in these techniques. The algorithm analyzes the input image then applies appropriate image processing techniques and identifies regions that potentially contain license plates.

Image Preprocessing Algorithm is responsible for enhancing the quality of the detected license plate regions. It involves techniques such as image filtering, noise removal (KR Soumya, A Babu, L Therattil, 2014) etc. These techniques help improve the contrast, eliminate noise, and ensure uniformity in license plate images which will help in accurate character segmentation and recognition.

Character Segmentation Algorithm focuses on extracting individual characters or symbols from the license plate region. It identifies the boundaries of each character, separates them from the license plate image, and prepares them for further processing.

Character Recognition Algorithm is responsible for interpreting and recognizing the separated characters. It may employ various techniques such as including optical character recognition (OCR) and machine learning algorithms The algorithm compares the segmented characters with known character templates or models to identify the alphanumeric information.

B. Design Analysis

A modular design approach can be adopted for the license plate recognition system. Each stage of the system, such as license plate detection, image preprocessing, character segmentation, and character recognition, can be implemented as separate modules. This design enables flexibility, scalability, and easy integration of additional features or enhancements in the future.

OpenCV libraries provide a wide range of functions and algorithms for computer vision and image processing. Designing the system to fully use the capabilities of OpenCV requires efficient and optimized implementation. The system can utilize OpenCV's image processing functions.

1) Workflow Analysis

The proposed workflow begins with acquiring input images or video frames containing vehicles and license plates.

License Plate Detection	The input images or frames are processed using the license plate detection algorithm to identify potential license plate regions.
Image Preprocessing:	The detected license plate regions undergo image preprocessing, where techniques such as filtering, thresholding, and morphological operations are applied to enhance the quality of the license plate images.
Character Segmentation:	The preprocessed license plate images are then analyzed using the character segmentation algorithm to separate individual characters or symbols.
Character Recognition:	The segmented characters are passed through the character recognition algorithm to interpret and recognize the alphanumeric information. This involves comparing the characters with known templates or classification models.
Post-processing and Output:	The recognized characters are post-processed to get the final output which includes the recognized license plate information. This can be used for various applications such as vehicle tracking, parking management, or law enforcement.

Fig. 4. Workflow matrix.

III. DESIGN AND METHODOLOGY

A. Research Paradigm

The development of computer vision has its roots in the mid-20th century, and it gained significant attention and improvement with the advancement of digital imaging and computing technologies. Initially, this technology was used primarily for printed text recognition but over time it has evolved (Rafael C. Gonzalez and Richard E. Woods, 2008).

The significance of optical character recognition lies in its ability to bridge the gap between the human and digital worlds (Tony F. Chan and Jianhong (Jackie) Shen, 2005). This technology is found in industries like finance, healthcare, legal, education, publishing, and more. It has streamlined processes by reducing manual data entry, enhancing accessibility for visually impaired individuals. I have decided to use OpenCV libraries in the development of this program. OpenCV is an open-source computer vision and image processing library that provides a wide range of functions and tools for handling images and videos. It was originally developed by Intel. OpenCV has become a widely used library in the field of computer vision due to its versatility, efficiency, and extensive documentation.

For the positivist research paradigm, a valid hypothesis is "A free OCR tool developed for character recognition of license plates are able to achieve a satisfactory level of accuracy which makes it a viable resource for the public to assist law enforcement in criminal investigations".

In order to carry on with the validity of the paradigm we can consider the independent variable to be the character recognition software that is being developed and as the dependent variable – the accuracy of character recognition. For the research design I propose a controlled experimental design. A dataset of license plate images is collected, containing various license plate types and fonts with different color backgrounds. We can divide the dataset into two groups called OCR group and manual recognition group (V Khare, P Shivakumara, CS Chan, T Lu, 2019). (Manual recognition meaning identifying characters with the naked eye) For the OCR group license plate images are processed using OCR software to extract characters and for the manual recognition group human participants manually transcribe characters from license plate images. Another experimental design we can use is: Collecting a diverse dataset of license plate images encompassing different license plate types, fonts and under various lighting conditions and angles. Then we can divide the dataset into two groups called the Free OCR group and the Professional OCR group. The free OCR tool is being developed using open-source resources and technologies available such as OpenCV and for the professional OCR tool we will use a commercial solution such as Apple's OCR available in their operating system as a benchmark. The license plate images are processed using both the free OCR tool and the professional OCR tool. The recognized characters are then compared with the actual characters on the license plates.

By this method the accuracy rates and error rates of both free and professional OCR tools can be calculated and compared. (G Bradski - Dr. Dobb's Journal, n.d.)

B. Research Approach

In this study, I have adopted an inductive research approach to investigate the factors influencing user preferences in utilizing Optical Character Recognition (OCR) tools for one-time License Plate Recognition (LPR) needs. The observations revealed a noticeable trend wherein individuals often resist paying substantial amounts for a single need of recognizing characters on a license

plate. The study aimed to understand this pattern and its implications, leading to the formulation of a theory suggesting that the availability of a privacy-conscious, publicly accessible free OCR tool could significantly influence users' choices and behaviors.

For this I published a survey where I collected data from 38 participants that represent a range of demographics and familiar with image processing applications and license plate character recognition programs. Participants were surveyed to gauge their awareness of LPR tools, willingness to pay for one-time LPR needs and their opinions on privacy concerns associated with OCR tools. The results show us that A significant percentage of participants specially those who are familiar with license plate character recognition programs expressed their reluctance to pay a considerable amount of money for a single use of an optical character recognition tool made for character recognition on a license plate.

We can draw a theory from this data. That if a publicly available free OCR tool for LPCR that doesn't store the user input images is likely to be utilized by the public instead of paying a considerable sum of money or subscribing to a monthly rental for the tool because the majority of the users will not have the need to use such a tool because most of the time they want to use a LPCR is when a crime is occurred such as grand theft auto or a hit and run etc. and they most probably suffered financial losses and tend not to spend the rest of it immediately. Also, if it is free, the bystanders will have no problem utilizing such a tool because it's free of charge.

Based on this information we can suggest that user preferences are heavily influenced by economic factors because in Sri Lankan culture most people value free or pirated versions (because that is normalized and even cause to humble brags in Sri Lanka) of software instead of purchasing them. The observed pattern of resistance to high costs aligns with broader trends of cost-conscious nature of most people. The proposed theory offers a potential explanation for the reluctance of users to opt for a paid tool when a privacy-conscious free alternative is available.

C. Research Strategy

This research project employs a case study approach as the research strategy because of its suitability in exploring real-world phenomena. For the methodology steps.

For sampling I collected data from 38 Participants to represent a diverse range of demographics and familiarity with LPCR technologies. Then the qualitative data is collected through surveys participants are asked about their awareness of LPR tools, willingness to pay for OCR services and concerns about their data privacy. Then I can use statistical analysis to determine their preferences.

With that data collection there is an evident pattern we can notice where a majority of the participants express reluctance to pay a minimum of $15 (roughly equal to Rs.5000) for one time character recognition event using an OCR tool. Therefore, we can arrive at the conclusion that economic considerations play a significant role in their decisions.

The case study approach allows for an exploration of these factors within a real-world context which provides us with a good understanding of the user's behavior. So, in conclusion of this chapter we could state that through this case study approach the research project offers insights into the complex interplay of user preferences, economic factors, and privacy concerns in the context of a license plate character recognition tool usage.

D. Fact Collection Mechanisms

Out of the fact collection mechanisms available for us the selection of questionnaires as the data collection mechanism for this research project holds several advantages that would align well with the research context. Questionnaires offer a structured approach to gathering insights from a diverse participant pool encompassing individuals with varying educational backgrounds and various degrees of familiarity with image processing tools and LPR technology. By utilizing questionnaires, we are able to efficiently capture respondents' preferences, behaviors and concerns.

We can use the quantitative nature of the questionnaire to identify the patterns and relationships among the participants.

When we study the responses, we can understand that a substantial proportion of participants demonstrated previous experience with image processing software while expressing a preference for free solutions. Moreover, the nuanced viewpoints regarding privacy concerns and acceptance of character recognition processes are integrated into the questionnaire.

I published the questionnaire on 17th of August and by 20th evening there are 39 responses which is a fair number of responses in order to get a general understanding.

The data collection process was initiated through questionnaires, with a total of 39 participants providing their responses. The participant pool demonstrated a diverse educational background, with 19 individuals holding bachelor's undergraduate degrees, 10 possessing a bachelor's graduate degree, 5 pursuing master's degrees, 3 being doctors, and 1 having other professional qualifications. This diversity ensured a well-educated respondent group.

Of the total respondents, 36 individuals acknowledged previous experience with image processing software, including applications. Remarkably, 29 participants confirmed that they had never purchased the premium versions of these applications, indicating a general inclination towards utilizing free software solutions.

Of the respondents familiar with image processing software, 7 individuals indicated their willingness to invest in premium versions of such tools.

Table 1. Resopnses

This detail signifies that a subset of the sample population is amenable to spending money on software products.

A noteworthy point is that only 8 out of the 38 respondents reported prior use of License Plate Recognition (LPR) tools. This finding suggests that the majority of participants were not acquainted with LPCR technology.

Interestingly, 24 participants stated that their encounters with license plates with the character 'shree' was almost never, whereas 8 respondents reported occasional occurrences of such plates that contain the letter 'shree'. This observation is particularly relevant considering the specific context of the Sri Lankan license plate systems historical use of special characters.

An intriguing aspect is that, despite acknowledging that the OCR process may not reveal vehicle ownership or other private details, 11 respondents expressed concerns about privacy violations, rejecting the idea of subjecting their license plates to an OCR process. Conversely, 24 participants indicated no qualms about their license plates undergoing OCR.

E. Research methodology Execution Workflow

This table outlines the key steps in the research methodology and execution workflow for developing a license plate character recognition application.

F. Design Methodology

SSAD method – Structured System Analysis and Design.

Description	Key Points
Problem identification	Lack of free license plate character recognition tools for one-time usage due to expensive subscriptions. High cost of existing solutions prevents one-time users from accessing this technology.
Relevance Justification	Providing a free solution can democratize access and promote innovation in related fields and help law enforcement agencies.
Define and Finalize Objectives	Develop a free, accurate, and efficient license plate character recognition application.
Data Management and Handling	Utilize open-source technologies which are developed on diverse datasets. Subject to preprocessing using image processing technologies.
Evaluation and communication	Evaluate the accuracy and efficiency of the application through testing on real-world license plate images. Use appropriate metrics (e.g., accuracy, precision, recall) to quantify the performance.

1) Feasibility study

Identify the need for a cost-effective license plate character recognition solution. Then assess the technical feasibility of using OpenCV for image processing. Analyze potential user scenarios where one-time character recognition is needed.

2) Requirement Analysis

Gather requirements from users where they had to utilize a LPCR but didn't want to spend a lot of money for a one-time requirement. Define input formats (image files), output requirements (recognized characters in plain text) and the constraints (like lighting conditions, noise etc.)

• System Design

• Design the architecture for the proposed solution.

3)Implementation

Develop the program. Utilize OpenCV for image manipulation, feature extraction and character recognition.

4)Testing

Test the program against real life images of license plates.

G. LESP Aspects

Discussing the legal, ethical, social and Professional aspects of this research.

Aspect	Domain	Fact

Legal Aspects	Privacy Laws	Ensure that the application complies with data privacy regulations, especially when dealing with images that may contain sensitive information like license plates.
	Intellectual Property	If I am planning to using third-party libraries, APIs, or algorithms, make sure I have the legal right to use them in the application.
Ethical Aspects	Data Usage	Handle user data ethically and transparently. Clearly communicate how data will be used and obtain consent from users if needed.
	Accuracy	Strive for accuracy in recognition to avoid potential misidentifications that could have serious consequences.
Social Aspects	Accessibility	Ensure that your application is accessible to a wide range of users.
	Being versatile	Consider the diverse needs of users from different backgrounds and regions when designing.
Professional Aspects	Adhere to best practices.	
	Honesty	Be transparent about the capabilities and limitations.

TABLE2. LESP EVALUATION MATRIX

IV. SYSTEM REQUIREMENTS SPECIFICATION

A. Stakeholder Analysis

Given below is the stakeholder analysis according to the Onion Model, starting from the innermost layer (core stakeholders) and moving outward.

1) Core Stakeholders (Inner Layer) - Project Owner

- Role: Sole developer, owner, and project manager.
- Interest: Project success

2) Stakeholder Layer 1 (Direct Stakeholders) End Users

- Individual Users: Seeking accurate character recognition for personal use.
- Law Enforcement Agencies: Interested in efficient license plate character recognition for investigations.
- Private Investigators: Need an effective tool for their investigative work.
- Security Companies: Looking for a dependable solution for surveillance and security purposes.

3) Website Visitors

- Role: General visitors seeking information about the application.
- Interest: Understanding the application's capabilities and benefits.

4) Stakeholder Layer 2 (Indirect Stakeholders)

- Other Software Developers: Monitoring the application's features and market position.
- Commercial Providers: Maintaining their market share against the cost-effective alternative.

5) Privacy Advocates

- Privacy-Conscious Users: Prioritizing data security and privacy.

6) Regulators

- Ensuring the application complies with data protection and privacy laws.

7) Stakeholder Layer 3 (External Stakeholders)

- Role: Hosting and infrastructure provider.
- Interest: Ensuring the reliability and security of their hosting services.

8) Stakeholder Layer 4 (Environmental Stakeholders)

- Independent Reviewers
- Role: Testing how accurate the system is against various conditions.
- Interest: Deeming whether the application is fit to use.

9) Technology Suppliers

- OpenCV and Tesseract Providers: Suppliers of the OpenCV and Tesseract libraries used in the application.
- Interest: Ensuring the reliability, updates, and support for the technology components integral to the application.
- Influence: The quality, performance, and compatibility of these technologies directly impact the application's functionality and success.

B. *Onion Model of the research*

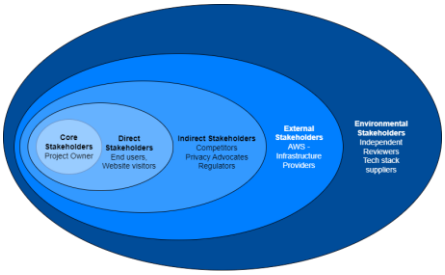


Fig. 5. Onion Model

C. *Operationalization Process*

1) Mapping data collection with objective results review

Understanding User Demographics and Preferences. Age: This question serves the purpose of profiling potential users by age, enabling a better understanding of the target age group for the free license plate character recognition software. The age distribution data indicates that a substantial portion, approximately 50%, falls within the 24-30 age range. This age group is a key demographic for the software's development and marketing efforts. Additionally, 35% of respondents are between 18-24, suggesting a younger user base, while 12.5% are in the 30-36 range, and 2.5% are between 36-45. Understanding the age demographics helps tailor user interface design and marketing strategies to appeal to the most relevant age groups.

Educational Qualifications: By collecting information on respondents' educational backgrounds, this question provides insights into potential users' technical proficiency and preferences. The majority of respondents, a significant 50%, hold BSc degrees as undergraduates, indicating a tech-

savvy user base. Additionally, 25% are graduates with bachelor's degrees in science, 15% are pursuing master's degrees, 7.5% have attained doctorates (MBBS or PhD), and 2.5% have other professional qualifications. This diverse educational background suggests that the software should cater to users with varying levels of technical expertise, highlighting the need for a user-friendly design.

Assessing Familiarity with Technology and Image Processing Applications Experience with Image Processing Applications: This question gauges respondents' prior experience with technology and image processing applications, which informs decisions related to user interface design and ease of use.

A significant percentage of respondents, 95%, have experience with image processing applications, indicating a high level of familiarity with the technology. This finding underscores the importance of designing the software with user-friendly features and interfaces, catering to users with prior experience in image processing. Determining the Need for a Free License Plate Character Recognition Software. Usage of Paid vs. Free Image Processing Applications: This question examines respondents' preferences for free software and their willingness to pay for similar services, validating the need for a free option.

The survey reveals that 75% of respondents have never felt the need to purchase a premium image processing application because free versions have sufficed for their needs. Furthermore, 17.5% have used paid versions, while 7.5% have never used such applications. The majority's strong inclination toward free options supports the objective of offering a free license plate character recognition software, especially during urgent or challenging situations.

In order to evaluate familiarity with License Plate Character Recognition Programs, Usage of License Plate Character Recognition Programs: This question assesses respondents' familiarity with license plate character recognition programs and their potential interest in using a free alternative. The survey data indicates that 80% of respondents have prior experience with license plate character recognition programs, suggesting a substantial potential user base for a free alternative. This familiarity positions the software to meet the needs of users who have already engaged with such technology.

Assessing Willingness to Use a Free License Plate Character Recognition Software. Willingness to Pay for One-Time Image Recognition: This question explores respondents' willingness to pay for image recognition, emphasizing the preference for a free option when available.

The survey responses reveal a strong preference for a free software solution, with 95% of respondents stating they would not pay for a one-time image recognition incident if a free option were available. Only 5% expressed willingness to pay under specific circumstances. This underscores the importance of offering a free license plate character recognition software to cater to the majority's preferences, especially in challenging situations, such as identifying a thief during a burglary. Understanding the Occurrence of Special Characters on License Plates. Frequency of Special Character License Plates: This question helps understand the frequency of encountering special characters on license plates, which may influence the software's character recognition capabilities.

The responses provide valuable insights into the occurrence of special character license plates. Approximately 65% of respondents indicated that they almost never see license

plates with special characters. Another 20% mentioned seeing such plates as a very rare occasion, more than a week apart. Only 15% reported seeing special character license plates less than twice a week. This data suggests that there may not be a high demand for a specialized character recognition dataset, as the majority of users do not frequently encounter special characters. Therefore, a generally available character recognition model may be suitable.

Addressing Privacy Concerns and Preferences. Attitudes Toward Vehicle Plate Image Usage: This question assesses respondents' privacy concerns regarding the use of their vehicle's plate image.

The majority of respondents, 66.7%, do not mind if someone uses an image of their vehicle's plate without permission to identify letters and numbers. However, the remaining 33.3% consider it a privacy violation. Given that a significant portion expressed privacy concerns, the decision to develop a desktop application, rather than a web application that stores client data, aligns with the preferences of the majority, and addresses potential privacy issues.

By mapping the questionnaire responses in this detailed analysis, we can clearly demonstrate how each question contributes to the research objectives and how the insights obtained from the survey guide the development and marketing strategies for the free license plate character recognition software.

D. Functional and Non Functional Requirements

1) Functional Requirements

Image Recognition Functionality: The software must have the capability to process various image formats, including JPG, JPEG, PNG, and BMP. It should accurately identify and recognize alphanumeric characters (letters and numbers) on the license plate.

Real-time Video Processing: The software should support real-time video processing, allowing users to analyze and recognize license plates in live camera feeds. It must have the ability to process video streams efficiently with minimal delay.

Simple User-Friendly Interface: The user interface should feature a straightforward design with clearly labeled buttons and controls. It should provide real-time feedback during image or video processing, indicating the progress and results of recognition.

Free to Use: The software should be accessible and usable by anyone without any cost or subscription fees. There should be no hidden charges, trial periods, or limitations on the number of uses.

Privacy Protection: User data and processed images must not be stored locally or transmitted to external servers or third parties. The software should offer a clear privacy policy that outlines its commitment to user data protection.

Offline Mode: Ensure that the software can operate in an offline mode, providing character recognition without requiring a continuous internet connection.

2) Non Functional Requirements

Accuracy and Speed: The character recognition accuracy should exceed 95%, ensuring reliable identification of license plate characters. Image and video processing should occur in real-time, with recognition results delivered within seconds.

Usability: The user interface should follow established usability principles, such as consistency, simplicity, and responsiveness. Users should be able to perform tasks

without the need for extensive training or technical expertise.

Privacy and Security: User data and processed information should be encrypted and securely stored on the user's device. The software should not collect or transmit any personally identifiable information (PII) or sensitive data.

Reliability: The software should exhibit high stability, with minimal instances of crashes, errors, or unexpected shutdowns. It should have automated error reporting and recovery mechanisms.

Performance: The software should efficiently handle large image files, video streams, and high-resolution images without significant performance degradation. It should have the capability to utilize multi-core processors for parallel processing when available.

V. Testing and Evaluation

A. Chapter Overview

In this chapter, we dive into the testing and evaluation process of the character recognition program designed for license plates. We begin by examining the program's accuracy, scrutinizing its ability to correctly identify characters on various license plates in real-world license plates which will be mostly Sri Lankan plates. Our investigation goes beyond mere accuracy as we assess the program's overall performance and scalability, shedding light on its efficiency and adaptability to handle varying workloads and environments. The discussion extends to functional testing, where we analyze the program's compliance with its intended functionalities, ensuring it meets the specified requirements. Additionally, we unravel the intricacies of the evaluation workflow, providing insight into the systematic approach employed to measure the program's success and identifying areas for improvement. Through this comprehensive examination, we aim to ascertain the reliability and practicality of the character recognition program for license plates, ultimately contributing to its refinement and optimization for real-world applications.

B. Initialization and Data Review

A number of license plates will be tested against the system to calculate accuracy and performance. Because the system emphasizes on Sri Lankan demography the majority of said license plates will be Sri Lankan plates for testing the maximum practicality.

Fig. 6. Sample of the test data used for accuracy testing – Sri Lankan styled and international styled plates.

C. Accuracy and Functionality Testing

1) Accuracy Testing

During the rigorous testing phase, a diverse set of 16 images, spanning various formats, including JPG, BMP, and PNG, underwent meticulous scrutiny to assess the character recognition capabilities of the program. The results proved highly promising, with 14 of the images returning accurate character identifications, effectively deciphering the characters adorning the license plates. This demonstrated the program's proficiency in handling diverse image formats, further enhancing its adaptability for real-world applications.

Perhaps even more noteworthy, there were no occurrences of false positives throughout the testing process. The absence of false

positives underscores the program's precision and its ability to discern genuine characters from background noise or image artifacts.

It's essential to highlight that the two instances where accuracy faltered provided unequivocal indicators of a false read. In these cases, the OCR output exhibited a collection of jumbled and incoherent characters, serving as a clear signal to the user that an erroneous recognition had occurred. This inherent fail-safe mechanism within the program offers transparency to users, reducing the likelihood of misleading or confusing results.

In conclusion, these test results exemplify the program's robust character recognition capabilities and underscore its rarity in producing false positives, further solidifying its position as a dependable tool for accurate and coherent character recognition on license plates in diverse image formats.

2) Functionality Testing

In the comprehensive examination of the program's functionality, I conducted a series of rigorous tests aimed at assessing its character recognition capabilities on images with the most commonly encountered formats, including JPG, PNG, and BMP. It's worth noting that JPG and JPEG formats, with their nearly identical characteristics, were examined collectively within the JPG format testing category.

The results of these functional tests emerged as a testament to the program's impressive performance. In each instance, regardless of the image format, the program consistently delivered accurate character recognition outcomes. This remarkable consistency across diverse formats underscores the program's versatility and adaptability to handle a broad spectrum of image types. The successful character recognition across JPG, PNG, and BMP formats signifies the program's ability to seamlessly process images, regardless of the user's choice or the source of the image.

This exceptional performance in functional testing serves as a strong indicator of the program's robust and reliable functionality. It not only assures end-users of dependable character recognition but also highlights its readiness to excel in real-world scenarios, where diverse image formats are commonly encountered. This adaptability and consistent accuracy reinforce the program's potential as a valuable and versatile tool for character recognition on license plates, positioning it as a trustworthy solution for a wide range of applications.

D. Testing Workflow

The testing workflow for the OCR License Plate Recognition (LPR) program is a structured and multi-phase process designed to ensure the program's accessibility and functionality. The workflow begins with a critical first phase: the accessibility test. This phase focuses on assessing the user's access to the program via a dedicated website. The website offers two distinct types of downloads – the OCR program itself and the essential .NET runtime setup, which users may require for smooth program execution. Both of these downloads are securely stored on prominent cloud platforms, AWS and Google Cloud Drive, utilizing content delivery networks to enhance accessibility and download speed.

Once the accessibility phase confirms the website's functionality and the download links' reliability, we transition to the second phase – the application run test. In this stage, the program's core functionality is scrutinized, specifically its ability to process images without encountering any errors. Upon confirming that the application executes seamlessly, the testing workflow advances to the third and crucial phase: functionality testing.

Within the functionality testing phase, two essential sub-phases are undertaken. The first is the image format test, wherein the program's capability to effectively recognize characters in a variety of image formats, including JPG, JPEG, PNG, and BMP, is thoroughly assessed. Remarkably, the program exhibits exceptional compatibility with all tested formats, further solidifying its adaptability.

The second sub-phase within functionality testing is the accuracy test, a critical evaluation of the program's character recognition precision. A total of 16 license plate images, encompassing Sri Lankan styled license plates and international/miscellaneous styled plates, were meticulously examined. Impressively, 14 out of these 16 images returned accurate results, resulting in an impressive 87.5% success rate. These findings affirm the OCR LPR program's outstanding accuracy and reliability in recognizing characters across diverse license plate types, reinforcing its potential as a trustworthy and dependable tool for a wide array of applications.

VI. Concluding Remarks

A. *Accomplishments of research objectives*

In developing this License Plate Character Recognition Desktop Application, the primary objectives were aimed to create a user-friendly, privacy-conscious, and cost-effective solution for character recognition from license plates in scenarios where individuals are in distress or require quick identification of license plate characters without resorting to potentially expensive and privacy-invading alternatives.

The key accomplishments in this regard include the following. Privacy-Centric Approach: The application's desktop nature ensures that user data remains on their local machines. There's no need to upload sensitive images to external servers, addressing privacy concerns.

Cost-Efficiency: By offering the program for free, we have made it accessible to a broad audience, including individuals who might be facing financial difficulties and cannot afford paid alternatives.

User-Friendly Interface: The program features a user-friendly interface that requires minimal technical expertise, making it accessible to a wide range of users.

Web Accessibility: Hosting the program on a website accessible through a web browser further enhances its accessibility. Users can easily download and use the application without complex installation procedures.

Security Measures: Hosting the website on AWS virtual machines and protecting it with AWS managed web application firewall ensures the security of the application and user data.

B. *Problem Encounters*

The development of the License Plate Character Recognition Desktop Application encountered several challenges such as ensuring OCR Accuracy: Achieving consistently high OCR accuracy for license plates under varying conditions, such as different fonts, lighting, and image qualities, was a significant challenge. User Interface Design was challenging because designing a user-friendly interface that caters to both technical and non-technical users required careful consideration and iterative design.

Integration Complexity means integrating OpenCV and Tesseract into a cohesive desktop application was a complex task which required thorough testing and debugging.

C. *Self-Reflection*

1) Ideology about the research carried out.

This application's research and development are in line with the philosophy of offering a public service that gives people power in difficult situations. The dedication to privacy and accessibility emphasizes the idea that technology should be used for the benefit of society.

The ideology regarding the research and development of the License Plate Character Recognition Desktop Application is rooted in a deep commitment to leveraging technology for the betterment of society while upholding essential values of privacy, accessibility, and social responsibility.

Privacy-Centric Approach: Central to this ideology is a staunch belief in safeguarding individual privacy. The decision to create a desktop application, where users retain control over their data, is driven by the principle that technology should empower individuals without compromising their privacy. By ensuring that sensitive images remain on users' local machines and are not uploaded to external servers, we respect users' rights to data privacy.

Accessibility and Affordability: A core tenet of this ideology is making technology accessible to all, especially those facing distressing situations. The application's free-to-use nature and absence of financial barriers align with the belief that individuals, irrespective of their financial circumstances, should have access to vital tools and resources when they need them most. This inclusivity ensures that even in times of trouble, individuals do not face additional financial burdens.

Community and Social Responsibility: The ideology recognizes the importance of community engagement and social responsibility. The decision to actively foster a user community and gather feedback underscores a commitment to continuous improvement and responsiveness to user needs. By empowering users to shape the application's evolution, we acknowledge that technology should serve the community's best interests.

Technical Excellence and Learning: This ideology places value on technical excellence and a commitment to learning. The challenges encountered during the development process, such as improving OCR accuracy and designing a user-friendly interface, are seen as opportunities for growth and innovation. The project's learning curves are embraced as a means of expanding expertise and problem-solving capabilities.

2) Benefits Gained

The development of this application has yielded several benefits, including technical skill enhancement meaning the project has enriched technical skills in image processing, OCR technology, and software development. Community Contribution meaning the application's availability for free has made a meaningful contribution to the community, potentially assisting individuals in distress.

3) Learning Curves

The project has been a significant learning experience, involving technical challenges where I had to overcome technical challenges in OCR, image processing, and software development and learning to design for usability and accessibility to cater to a diverse user base.

D. Business Insights of the data

The License Plate Character Recognition Desktop Application, being a free-to-use program, presents a unique opportunity for sustainable monetization through the hosting website. The decision to monetize the website using advertisements is a strategic approach to support the ongoing development and maintenance of the

application while ensuring it remains accessible to all users.

Implementing ads on the website can generate a steady stream of revenue. By partnering with advertising networks or platforms, such as Google AdSense, The ability to display relevant ads to website visitors. The revenue generated from these ads can help cover operational costs, server hosting fees, and future development efforts.

The primary benefit of using ads as a revenue source is that it allows me to keep the License Plate Character Recognition Desktop Application free for users. This aligns with the original vision of providing accessible, cost-free assistance to those in need, such as homeowners in distress.

Ad-based monetization ensures a consistent and sustainable source of income. This revenue can be reinvested in the application, enabling ongoing improvements, updates, and enhancements. Users are not required to pay for the application or incur any hidden costs.

Instead, the program remains entirely free while ads generate revenue in the background.

1) Real-world applications

Law Enforcement and Public Safety: Police departments can use the application to identify and track vehicles involved in criminal activities based on license plate recognition. This aids in investigations, monitoring wanted vehicles, and enhancing overall public safety.

Private Security and Access Control: Private security firms and gated communities can employ the technology for access control. Recognizing authorized license plates ensures only permitted vehicles enter restricted areas, enhancing security.

Parking Management: Parking facilities and services can utilize license plate recognition for ticketless parking, efficient entry/exit, and tracking parking duration, improving user convenience and revenue collection. Toll Collection and Express Lanes: Toll booths and express lanes on highways can automate toll collection through license plate recognition, reducing congestion and improving traffic flow.

Traffic Management and Regulation Enforcement: Traffic authorities can monitor and enforce traffic regulations, such as identifying vehicles with expired registrations, tracking traffic violations, and optimizing traffic flow.

Vehicle Inventory Management so that car dealerships and rental agencies can streamline vehicle inventory management by quickly identifying cars based on their license plates, facilitating tracking and organization. And logistics and supply chain companies can use the technology for tracking and managing shipments, ensuring efficient transportation and timely deliveries. And automated entry systems to secure facilities like corporate campuses, airports, and educational institutions can automate entry and exit through recognition of authorized vehicles' license plates.

These real-world applications highlight the versatility and significance of the License Plate Character Recognition Desktop Application. Its potential to enhance security, streamline operations, and improve safety spans across various industries and sectors, making it a valuable tool with wide-ranging benefits.

REFERENCES

[1] F Xie, M Zhang, J Zhao, J Yang, Y Liu, 2018. A robust license plate detection and character recognition algorithm based on a combined feature extraction model and BPNN, s.l.: s.n.

[2] G Bradski - Dr. Dobb's Journal, n.d. The openCV library., s.l.: s.n.

[3] KR Soumya, A Babu, L Therattil, 2014. License plate detection and character recognition using contour analysis, s.l.: s.n.

[4] L Xiaobo, L Xiaojing, H Wei, n.d. Vehicle license plate character recognition, s.l.: IEEE.

[5] M Egmont-Petersen, D de Ridder, H Handels, n.d. Image processing with neural networks—a review, s.l.: s.n.

[6] M Sonka, V Hlavac, R Boyle, 2013. Image processing, analysis and machine vision, s.l.: s.n.

[7] M van Heel. G Harauz, EV Orlova, R Schmidt, n.d. A new generation of the IMAGIC image processing system. s.l.:s.n.

[8] MMP Petrou, C Petrou, 2010. Image processing: the fundamentals. s.l.:s.n.

[9] Rafael C. Gonzalez and Richard E. Woods, 2008. Digital Image Processing, s.l.: s.n.

[10] SL Chang, LS Chen, YC Chung, n.d. KR Soumya, A Babu, L Therattil, s.l.: s.n.

[11] T Acharya, AK Ray, 2005. Image processing: principles and applications, s.l.: s.n.

[12] T Pavlidis, 2012. Algorithms for graphics and image processing, s.l.: s.n.

[13] Tony F. Chan and Jianhong (Jackie) Shen, 2005. Image Processing and Analysis: Variational, PDE, Wavelet, and Stochastic Methods, s.l.: s.n.

[14] TS Huang, WF Schreiber, OJ Tretiak, n.d. Image processing: the fundamentals, s.l.: Proceedings of the IEEE.

[15] V Khare, P Shivakumara, CS Chan, T Lu, 2019. A novel character segmentation-reconstruction approach for license plate recognition, s.l.: s.n.

[16] YJ Zhang, 2017. Image processing, s.l.: s.n.

[17] Documentation, O., n.d. [Online] Available at: https://docs.opencv.org/4.x/

[18] Joao Manuel R.S. Tavares, 2010. Image Processing and Analysis: Applications and Trends, s.l.: University of Porto.

[19] Manjit Sandhu, 2022. An Application of Image Processing, s.l.: Indian Institute of Technology Ropar.

[20] Tartu, U. o., n.d. [Online] Available at: https://sisu.ut.ee/imageprocessing/book/1

[21] Surinder Kaur, 2021. Vehicle License Plate Recognition, Bharati Vidyapeeth College of Engineering

[22] OpenCV Documentation – Cascade Classifier [online]

[23] https://www.sciencedirect.com/topics/engineering/image-processing

Design and Development of a Mobile Application for Facilitating Donations to Elder Care Homes

Nisansala Kaluthanthri
Faculty of Computing
NSBM Green University
Pitipana, Homagama
nisansalasanda1999@gmail.com

Mohamed Sapraz
Faculty of Computing
NSBM Green University
Pitipana, Homagama
shafraz@nsbm.ac.lk

Abstract*—This study addresses the lack of a unified donation platform for elder care homes in Sri Lanka, aiming to develop a user friendly mobile application. The central goal is to streamline and simplify the donation process by understanding user preferences, challenges faced by donors and elder homes, and the intricacies of donation mechanisms. Key objectives include raising awareness, ensuring transparency, and catering to the diverse needs of donors and elder homes. The research methodology integrates qualitative analyses to comprehensively grasp donation behavior and societal perceptions. The data was gathered through a questionnaire and after an extensive review of findings through the thematic analysis approach 4 Themes and 10 Sub Themes were identified. The proposed mobile application designed based on the findings expands the scope of donations to essential items, fostering a culture of giving and shared responsibility in society, with the vision of initiating a movement where care knows no boundaries.

***Keywords*—Elder Care Homes, Mobile Application, Donations, Transparency, User Preferences, Social Welfare**

I. INTRODUCTION

Sri Lanka is undergoing a period of accelerated aging, leading to an increased demand for long-term care services. In 2012, 12.4% of Sri Lanka's population was over 60, and it's projected to reach 21.9% in 2031 and 27.4% in 2050[1]. Changing family dynamics have left older individuals without secure and easily accessible amenities at home[2]. Elderly homes, provided by a mix of for-profit, government, and charity organizations, have traditionally been the conventional solution to address the rising elderly population in the country. The growing elderly population has led to increased demand for elder care homes, with approximately 255 homes serving around 7,100 elders in Sri Lanka.

Elder care homes in Sri Lanka offer services like caregiver training, long-term care, infrastructural improvements, and active aging advocacy. These facilities heavily rely on donations, whether monetary or in-kind, facilitated through online platforms or NGOs. However, they face challenges such as inadequate funding, staffing issues, and concerns about the quality of care, leading to issues of abuse and public apprehension regarding the well-being of elders and donation processes across multiple platforms.

Online donations have become a convenient method for charities, benefiting both donors and recipients[3]. While online platforms enable recipients to publish donation requests and information, many elder care homes in Sri Lanka lack a digital strategy. This absence hinders their ability to reach potential donors in a competitive online landscape where other organizations have a more focused online presence.

People hesitate to donate for various reasons, including financial constraints, distrust in how charitable organizations utilize funds, a preference for donating time or goods over money, the belief that their contribution may be insufficient, uncertainty about transparency in organizations, and a lack of convenient online platforms.

Complex donation processes and concerns about loyalty to charities also contribute to hesitancy[4]. Another significant factor for potential donor being hesitant in making donations would be lack of proper education about charitable education and their environments[5]

This study addresses the deficiency in the online donations sector by identifying a lack of a unified platform for all stakeholders to discuss and devise solutions for simplifying the donation process. The gap in community understanding about elder homes and the absence of a comprehensive mobile or web application consolidating information on all elder homes are key issues that the proposed system aims to resolve. The goal is to create an efficient and effective solution that enables seamless contributions and access to relevant details for donors.

II. Literature Review

In this section existing systems will be reviewed in the field of online donations.

A. I-Do

I-Donate is a mobile app designed for online fundraising, catering to various organizations like aid groups, children's homes, and elder homes to secure financial support from the community for their ongoing operations. However, the app has encountered challenges, including design issues in user interfaces that led to confusion among users. Additionally, its development is not yet complete, attributed to a limited user base for gathering feedback and identifying potential drawbacks and issues in the application[6].

B. E-Donation

The E-Donation is a mobile app originating from Pakistan, created to facilitate the sharing of surplus food from events, restaurants, and volunteers. The platform enables all users to create accounts, allowing donors to share their excess food by responding to requests made by NGOs[7].

C. UNNATI SAMAJ

A web-based application designed to streamline diverse donations, including food, clothing, and essential items, enables donors to contribute based on their capacity to the nearest NGOs without the need for additional efforts in specifically searching for nearby organizations[8].

D. Helping Hands

"Helping Hands" is a mobile application developed to streamline charitable donations to orphanages and old age homes in India. It tackles the difficulty of identifying suitable non-profit organizations by providing a thorough list for donors to select from. With a layered architecture ensuring a strong user experience, the app connects donors with non-profit organizations, simplifying the donation process and benefiting both parties [9]

Examining the current literature on donation platforms reveals the absence of a unified online platform in Sri Lanka that allows donors to access information and contribute to multiple elder homes simultaneously. Additionally, it reveals a trend where existing platforms often concentrate on specific donation aspects (such as food or clothes) and lack traceability options for donors to track their contributions. Consequently, this research seeks to address these shortcomings by proposing a technological solution through the development of a mobile application for donations to elder care homes.

III. Methodology

This research project adopts the Design Science Research Methodology, an approach aimed at generating innovative solutions for practical issues. It involves a systematic process of creating effective solutions, evaluating their success, and communicating the outcomes in addressing real-world problems [10]

A. Data Collection Mechanism

The research project employed a qualitative data collection method through the use of a questionnaire. The questionnaire

was meticulously crafted to extract crucial insights aligned with the fundamental research objectives pertaining to donors. The survey was administered to a total of 30 participants and the responses encompassed a diverse spectrum of ages, provinces, employment statuses, and genders, thereby ensuring the compilation of a comprehensive and varied dataset.

B. Data Analysis

In this research, thematic analysis was employed as a systematic approach to identify, organize, and interpret themes within the qualitative data. The initial step involved a thorough review of raw data extracted from questionnaire responses to gain a profound understanding of participants' perspectives. Initial codes were generated during this phase, representing recurring ideas and concepts within the dataset. Subsequently, these codes were systematically grouped into overarching 4 main themes that encapsulated the essence of participants' narratives. Further refinement led to the creation of 10 subthemes, enabling a nuanced exploration of the dataset.

Themes	Sub Themes
Altruistic Motivations	Personal Values Social Influence Recognition
Trust and Transparency	Trustworthiness Transparent fund usage
Convenience and Accessibility	Easy Navigation Secure Payment Unified Donation Platform
Information Preferences	Detailed Information Real Time Updates

Table 1. Themes and Sub Themes

C. Thematic Analysis Results

The thematic analysis profoundly influenced the design of a mobile app aimed at facilitating donations to elder care homes in Sri Lanka. Key themes such as altruism were translated into the app through impactful stories and testimonials, fostering emotional connections with users. Social influence was leveraged via a social feed feature, enhancing community engagement by enabling users to share donation requirements. Additionally, recognition was prioritized through features like a "Donor Wall" ensuring donors feel valued and appreciated. Trust and transparency were embedded through detailed profiles of elder care homes and transparent fund usage tracking. The app prioritizes convenience with minimalist design and secure payment gateways. It also serves as a unified platform for exploring elder care homes, offering detailed project information, interactive infographics, real-time updates, and notifications to enhance user engagement and trust, ultimately encouraging regular contributions to elder care homes in SriLanka.

IV. Proposed System

Once a new user registers with the system as a Donor they'll be able to view the home page of the mobile app where they can mainly view donation categories, emergency requirements posted by elder care homes and a list of elder care homes registered within the application. Also, once they select an emergency requirement posted the user will be directed to the interface as displayed in Fig 1 given below.

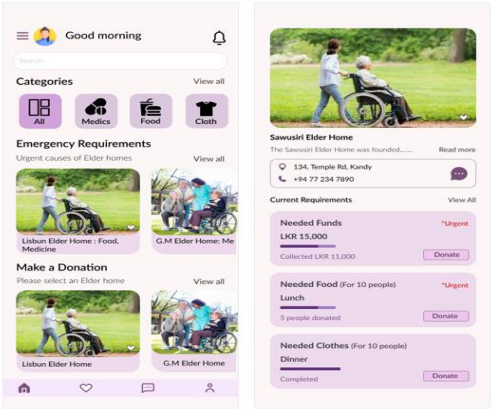


Fig 1. Home page and Elder Home Emergency requirement page

Once a donor decides to make a fund transfer, they'll encounter the following process displayed in Fig.2 given below.

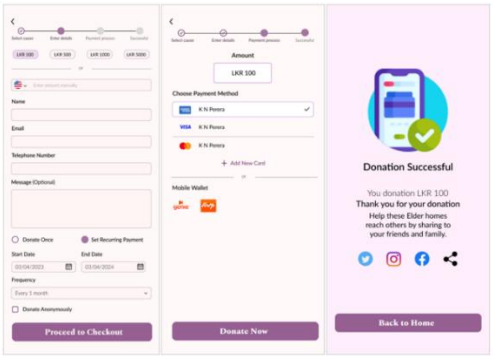


Fig 2. Monetary donation process interface

If the donation is related to a cloth or food items, then the user will have the ability to track the donation until it's received by the respective elder care home as shown in the Fig 3 given below.

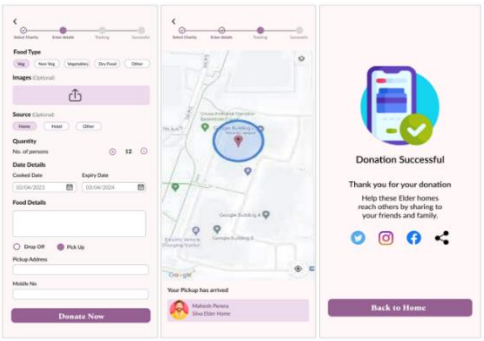


Fig 3. Food donation process interface

The proposed mobile application also provides the option for users to view all the donations they have made and a summary of the donations by categories once they select the Donations option in the User profile and also to view the "Donor Wall" feature in the community section which displays top donors of the application within a specific month as displayed in the Fig 4 given below.

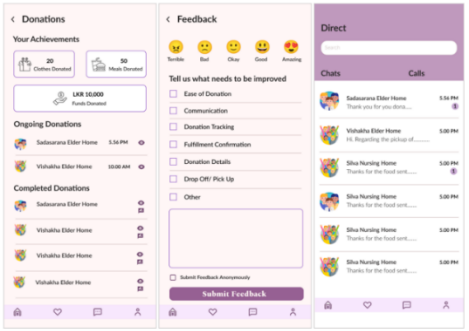


Fig 4. Profile page, Donations page and Community page

V. Future Work

Looking forward, it is advisable to concentrate on the development of this proposed mobile application, tackling the challenges identified in the research phase. Collaborating with technology specialists, user experience designers, and social workers could enhance the project. Furthermore, implementing pilot studies involving donors and elder care homes which would offer valuable real-world feedback crucial for refining the application.

VI. Conclusion

The research expedition highlighted the significance of combining empathy with technology in the realm of social welfare. It underscored the effectiveness of community-driven initiatives and their positive influence on the well-being of vulnerable individuals. The research yielded valuable insights into the specific needs of both donors and elder care homes. The envisioned mobile application has substantial real-world potential, capable of fostering community engagement and philanthropy once fully developed. Beyond its initial focus on elder care, the application's framework could be adapted to support diverse charitable causes, magnifying its impact and establishing a network of socially conscious individuals.

Acknowledgment

I express profound gratitude to my esteemed research supervisor, Dr. Mohamed Sapraz, for his steadfast guidance, limitless patience, and valuable feedback for this research endeavor. I also acknowledge and

appreciate NSBM Green University for providing the necessary resources that fueled this academic journey. My heartfelt thanks extend to friends and family whose willingness to dedicate time for surveys, data, and thoughtful insights played a pivotal role in shaping the success of this project.

REFERENCES

[1] S. Lanka, 'Data Collection Survey on Aging Sector in Sri Lanka Final Report Data Collection Survey on Aging Sector in Sri Lanka Final Report Sri Lanka Democratic Socialist Republic Aging Sector Information Gathering and Confirmation Survey CONTENTS', 2021.

[2] R. G. Ariyawansa, M. A. N. R. Perera, and H. E. S. Priyanka, 'Suitable Home Environment for Elderly People in Colombo, Sri Lanka', *Journal of Aging and Environment*, vol. 37, no. 1, pp. 85–99, 2023, doi: 10.1080/26892618.2021.2001708..

[3] T. Hou, K. Hou, X. Wang, and X. (Robert) Luo, 'Why I give money to unknown people? An investigation of online donation and forwarding intention', *Electron Commer Res Appl*, vol. 47, May 2021, doi: 10.1016/j.elerap.2021.101055.

[4] B. A. Sargeant, 'Building Donor Loyalty: The Antecedents and Role of Commitment in the Context of Charity Giving', 2007.

[5] X. Xu and Z. Li, "Literature Review on Elderly Charitable Donations," International Conference on Social Science, Education Management and Sports Education (SSEMSE 2015), Jan. 2015, doi: 10.2991/ssemse-15.2015.564.

[6] C. Z. Xuan, S. Y. Jing, J. C. L. Hin, Y. S. Qi, A. Z. Saidin, and S. N. Abdul Salam, 'IDO: USER EVALUATION OF ONLINE FUNDRAISING MOBILE APPLICATION', *Journal of Information System and Technology Management*, vol. 7, no. 29, pp. 186–194, Dec. 2022, doi: 10.35631/jistm.729017.

[7] M. J. Iqbal, 'E-Donation Mobile Application Automatic Brain Tumour Segmentation in Magnetic Resonance Images View project Brain Tumor Segmentation in MRI View project', doi: 10.13140/RG.2.2.22320.84483.

[8] S. Belekar, R. Rajput, K. Gharat, and P. Raut, 'MOBILE APPLICATION FOR DONATION OF ITEMS', 2021. [Online]. Available: www.viva-technology.org/New/IJRI

[9] S. Sai Chaithanya Elapanti and N. Sai Pinthepu, 'Pinthepu Nikhil Sai', *International Journal of Advance Research*, 2018, doi: 10.5121/csit.

[10] K. Peffers, T. Tuunanen, M. A. Rothenberger, and S. Chatterjee, 'A design science research methodology for information systems research', *Journal of Management Information Systems*, vol. 24, no. 3, pp. 45–77, Dec. 2007, doi: 10.2753/MIS0742-1222240302.

Stock Price Prediction Approach Using Machine Learning, Reinforcement Learning, And Sentiment Analysis of Stock News Headlines

Shemeen Fernando
Faculty of Computing
NSBM Green University
Homagama, Sri Lanka
fdoshemeen61@gmail.com

Rasika Ranaweera
Faculty of Computing
NSBM Green University
Homagama, Sri Lanka
ranaweera.r@nsbm.ac.lk

Dulanjali Wijesekara
Faculty of Computing,
NSBM Green University
Homagama, Sri Lanka
dulanjali.w@nsbm.ac.lk

Abstract*—The role that a stock market plays within a financial market is indispensable. The stock market can be identified as a place where we can trade stocks of companies among the participants of the market for an agreed price. When investing in the stock market, investors must understand the price fluctuations of various market parameters. This is the stage where the people who are new to the stock market get stuck due to lack of price fluctuation awareness. Traders focus on finding the best trading positions for several reasons. To address this, this research proposes a stock price prediction system that integrates machine learning, reinforcement learning, and sentiment analysis using Natural Language Processing (NLP) on news articles related to specific stock symbols. The goal is to reduce trading risk and maximize profits. Within this research paper my goal is to propose a stock price prediction system by using machine learning and reinforcement learning approaches along with stock news sentiment analysis to mitigate the risk in stock trading and also in maximizing the profits over the existing methods. Since predicting the exact future price of a stock is impossible, machine learning and deep reinforcement learning approaches can be used to predict the stock market trends to generate buy and sell orders within the market according to the predicted price. In this paper it proposes an ensemble stock price prediction model, combining three machine and deep learning models which are LSTM, ARIMA, and reinforcement learning to forecast stock prices and the LSTM model includes sentiment analysis of daily stock news headlines as well. These models make individual predictions, and the final forecast is a weighted average by combining all the predictions from three separate models. The accuracy of the model is evaluated by comparing predicted prices with actual prices over a 7-day period. Overall, this project shows how combining different technologies can help us understand and navigate the complexities of the stock market

Keywords*—Stock Prediction, Machine Learning, Reinforcement Learning, Long Short-Term Memory (LSTM), Autoregressive Integrated Moving Average (ARIMA), Natural Language Processing (NLP), Sentiment Analysis

I. INTRODUCTION

In this research endeavor, the focus revolves around the development of an automated stock price prediction system, leveraging the power of information technology and machine learning techniques. The financial market's quest for effective prediction models has intensified with recent advancements, leading to the rise of algorithmic or quantitative trading. The project aims to bridge gaps in existing research by incorporating sentiment analysis from news platforms into the automated prediction system, acknowledging the

importance of external factors in trading automation. The overarching goal is to create a robust and accurate model that not only enhances decision-making but also addresses the challenges posed by emotional biases in manual predictions.

Research on the stock market predictability has a long history of financial economics and among all the methods which were used in stock market prediction such as statistical or econometric methods which are based on the analysis of the movements of the past market have been the most widely adopted methods [1]. Successfully predicting financial data has always been an interesting topic in the financial industry [2]. Numerous studies that were published are based on Machine Learning models which have relatively better performance than the already available classical forecasting techniques. As a result, the literature in the fields of finance and computational intelligence is constantly flooded with new publications and implementations.

Some of the previous research on stock prediction used past data [3][4][5][6][7], social media data [8][9][10][11][12] or news data [13][14][15][16] in-order to predict the stock market by using machine learning algorithms. In here, different predictive models have been suggested or proposed that use one or the other type of data. At the end these systems provide useful information to investors to make their investment decisions for buying, selling or holding a stock. But using one kind of data may not give increased prediction accuracy for a stock market.

Historical data has been used in a technical analysis approach in which mathematics is applied to analyze data for finding future stock market trends [13]. Researchers used different machine learning techniques, such as deep learning [16] and regression analysis [17], on stock historical price data, but it is important to include external factors because unexpected events expressed on social media and financial news can also affect stock prices. Some researchers like Vijh characterize stock markets as unpredictable, dynamic and non-linear in nature [18]. Because of the inherent uncertainty in market movement, stock market prediction is seen as a difficult undertaking when predicting financial data. The expectations of the traders, general economic conditions, political events, and other factors all have an impact on the stock market. As a result, forecasting stock values is frequently challenging.

The stock market is considered as a vital component of a country's economy [19]. The stock market is one of the biggest investment opportunities available to businesses and investors. It is also seen to be a favorable moment for investors to buy new stocks and make more money from dividends which are paid in the company's shareholder business program. An investor can also trade stocks as traders within the stock market. However, investing in the stock market is risky, but when you approach it in a disciplined manner, it simply could be one of the most efficient ways to enjoy substantial profits. Before buying stocks of a particular company, the first thing that investors do is to evaluate the company to decide before buying risky stocks. This study could entail looking at a company's performance on past data and financial news websites, among other things. Investors may be unable to analyze such a large volume of data. As a result, a decision support system that automates decisions is required.

II. Problem Statement

In the dynamic realm of the stock market, participants are engaged in a constant pursuit of lucrative trades, aiming to comprehend and capitalize on market price fluctuations. Traders strategically seek optimal positions to both maximize profits and mitigate risks, incorporating calculated risk-taking strategies, setting stop loss orders, and conducting comprehensive analyses of various factors, including news events, market patterns, and technical indicators. Success in this environment necessitates continuous adaptation and refinement of trading strategies over time. Recognizing the complexities of the ever-changing stock market landscape, numerous professional traders have turned to sophisticated algorithmic tools to navigate and thrive. In essence, the imperative for an automated stock price prediction system arises as a crucial need in facilitating well-informed trading decisions. Such a system is envisioned to empower traders by enabling them to discern optimal trading opportunities, minimize risks, and optimize profitability in the fast-paced and dynamic stock market environment.

III. Related Work

In order to create prediction models, stock market researchers have mined historical, news, and social media data using a variety of reinforcement learning and machine learning techniques. Before financial news platforms and social networking platforms were so common, historical stock price data was used in predicting future stock prices. Successfully predicting the future stock prices has some interesting benefits which usually affects the decision making of financial traders on purchasing or selling financial instruments. A study which was carried out with the aim of predicting the closing price of iShares MSCI UK, have used all together four data mining techniques which are artificial neural networks, SVM (Support Vector Machine), RF (Random Forest) and LSTM (Long Short-Term Memory networks) s and compared the daily closing prices of iShare MSCI UK starting from January 2015 to June 2018 [20]. Results of this study show that recurrent neural network method with LSTM block functions outperforms other methods.

Khan has developed a framework for predicting stock market trends using social media and news data as external factors by examining the effect of social media and financial news content in order to predict the stock prices for upcoming 10 days [21]. As a result of this study, they have found that social media and financial news content has a greater impact or an influence on stock prices. [22] recommended and used the basic approaches for stock price prediction and concentrated on the sentiment analysis model by applying various classification machine learning techniques and for a thorough strategy on leveraging news content and social media data along with historical and previous stock data. They have also analyzed the impact on news content and social media data on stock prices for a certain period of time. They suggest that polarity detection can be used in order to identify whether a certain news content as positive or negative and when a positive news effect is likely to increase the share market values and if its negative, then the price to decrease.

Some researchers have investigated the effect of public sentiments using twitter as the social media platform to analyze the correlation between them when predicting the stock prices of the tech companies which are the most

popular according to yahoo finance [23]. Their results show the effect of change in the stock market based on public sentiment which means the high chance of predicting the stock market according to them. But their analysis has not considered many factors which could affect stock prices. Several approaches have been presented to predict stock value movement patterns on a weekly forecast horizon using eight machine learning and four LSTM based deep learning regression models. Out of all the deep learning and machine learning-based regression models, it was discovered that the LSTM-based models' performances outperformed those of the prediction models based on machine learning [24].

The study of [25] states that they have tested hypothesis that sentiment analysis that they have done using Twitter data could provide additional information and increase the accuracy in predicting stock prices. A novel and a deep reinforcement learning approach for automating the stock transactions was carried out by [26] and proved the practicality of DRL when dealing with financial strategy issues by using a comparison between three classical DRL models which are DQN (Deep Q Network), DDQN (Double Deep Q Network), Dueling DDQN (Dueling Double Deep Q Network). It states that these three algorithms have a better performance and intelligence over the traditional transactions since they could quickly respond and adapt to changing market conditions. After comparing the results of three algorithms, it was concluded that DQN maximizes the decision benefits among the other two algorithms.

Some researchers like [27][28] have used ARIMA which is a regression model in forecasting financial data such as stock prices. This method of ARIMA can also be identified as a statistical method which is mostly used for decomposing and also in predicting stock price data. According to some researchers, there are many advantages in using ARIMA regression method in predicting stock prices, especially for short-term. SVM can be used in conjunction with the ARIMA model to anticipate stock prices, which are nonlinear in nature. The ARIMA model has been widely used for linear series, however it has been discovered that the model cannot readily handle nonlinear series [29]. The average global temperature time series data has also been predicted using the Basic, Trend Based, and Wavelet Based ARIMA Model variations; the Trend Based ARIMA Model outperforms the other variants in this regard [30].

When considering all the above information, researchers have utilized various machine learning and reinforcement learning techniques in order to predict future stock prices by analyzing historical data, news content and social media data. Studies have shown the effectiveness of approaches such as artificial neural networks, support vector machines, random forests and LSTM networks in predicting stock prices. incorporating sentiment analysis from social media and news content has been found to have a significant impact on predicting stock prices. Overall, these studies highlight the value of integrating data-driven techniques with sentiment analysis and deep learning to improve stock market predictions and informed trading strategies.

IV. Methodology

The choice of model technologies for building the automated stock price prediction system was

meticulously chosen by insights from the literature review. networks, known for their adeptness in capturing sequential patterns and long-term dependencies of time-series data such as stock price data, were selected to model and forecast stock prices based on historical data. The classical time series forecasting model, ARIMA, complemented LSTM in predicting stock prices, particularly when dealing with stationary time series data. To optimize stock trading decisions by predicting stock prices, Q-Learning which is a reinforcement learning technique, was integrated to provide valuable insights into buy/sell choices.

The methodology involved several key steps in creating the prediction framework. It began with clearly defining the research problem, focusing on identifying trends and predicting daily close prices for selected stock symbols which in the case of this research are AAPL, MSFT and ORCL. Data collection included extracting historical stock price using the Yahoo Finance API and historical news data using Alphavantage API. Preprocessing steps addressed data cleanliness, handling missing values, and removing outliers. Feature engineering incorporated factors like moving averages and technical indicators and model selection involved LSTM and ARIMA for predictions, along with a reinforcement learning model, specifically a deep Q learning agent. Sentiment analysis results were integrated into the LSTM model. Real-time data collection, user interface development, model evaluation, continuous optimization, and user support were integral components of the comprehensive methodology, ensuring the creation of a robust stock price prediction framework.

A. LSTM Model

The model's construction involves a meticulously designed stacked LSTM architecture with 4 LSTM layers, incorporating features such as closing prices, trading volumes, and sentiment scores. The training process includes careful scaling using MinMaxScaler and thoughtful transformation of historical data into sequences suitable for training, ensuring a comprehensive and effective approach.

B. ARIMA Model

The Autoregressive Integrated Moving Average (ARIMA) model, which is like a statistical powerhouse in time series forecasting, plays a pivotal role in capturing underlying patterns in sequential data like stock prices. Starting with meticulous data preprocessing, the ARIMA model is tailored for each ticker symbol. Careful parameterization with (p, d, q) ensures adaptability to different time series data, while the training process involves the creation of a robust and versatile model for future predictions.

C. RL Agent Model

Reinforcement Learning (RL) agents introduce intelligence into stock price prediction by learning optimal trading decisions through trial and error. In this model, a Q-learning-based RL agent is implemented by creating a dynamic gym-like environment for accurate stock price predictions. The training process, spanning episodes of historical data, results in the development of individualized RL agents for each stock symbol, showcasing adaptability to the unique characteristics of each asset.

D. Sentiment Analysis

Sentiment analysis emerges as a key component in deciphering the collective sentiment around specific stocks. Through

an automated process, textual data from news articles undergoes sentiment analysis, categorizing sentiments into positive or negative. This insightful analysis allows for the identification of articles likely to impact stock prices, contributing to a focused and data-driven approach in stock trading and investment.

E. Ensemble Model

The ensemble model ingeniously combines the predictions from LSTM, ARIMA, and RL models through a weighted average technique. Strategic weight distribution, considering the Root Mean Square Error (RMSE), ensures a nuanced approach where models with lower RMSE carry higher weights. The final output represents a unified prediction, harnessing the strengths of each model for accurate and robust stock price predictions. Table 1 illustrates the weight distribution of models based on their RMSE.

TABLE I. WEIGHT DISTRIBUTION OF MODELS

Stock Symbol	***Model***	***RMSE***	***Weight***
AAPL	ARIMA LSTM RL Agent	2.95 3.39 13.39	0.476 0.417 0.106
MSFT	ARIMA LSTM RL Agent	5.40 6.63 11.73	0.439 0.358 0.203
ORCL	ARIMA LSTM RL Agent	1.83 2.42 8.20	0.505 0.381 0.113

V. EVALUATION AND RESULTS

The rigorous accuracy testing and evaluation process of this model aim to thoroughly assess the reliability and precision of the predictive models. The test plan entails a daily evaluation over a one-week period, where each day involves utilizing the models to predict stock prices based on user-specified inputs. The predictions are then compared against the actual stock prices for the subsequent day, and the accuracy of each daily prediction is quantified by calculating the percentage of correctly predicted stock prices. This daily accuracy assessment provides valuable insights into the models' ability to forecast stock prices accurately. Over the entire one-week testing period, the daily accuracy percentages are averaged to derive an overall assessment of the models' performance. This comprehensive approach ensures that the models are evaluated consistently and for an extended duration, capturing the nuances of real-world stock price fluctuations. The methodology considers variations in stock prices and provides a robust measure of the models' effectiveness, contributing to the project's credibility and practical relevance.

The testing accuracy is measured individually for each of the three models - ARIMA, LSTM, and the RL agent - as well as for the ensemble model. This involves predicting stock prices for the upcoming 7 days across three different stock symbols (AAPL, MSFT, and ORCL) and comparing these predictions with the actual market prices separately. The percentage accuracy is then calculated for each model, reflecting how closely the predictions align with the real stock prices. This accuracy measurement is conducted across all three stock symbols, providing a comprehensive view of each model's performance across diverse financial instruments in real-world scenarios. The calculated accuracies, utilizing the specified equation, contribute valuable insights into the effectiveness and reliability of each model and the ensemble approach, enriching the research

findings with robust empirical evidence. Equation (1) shows how the accuracies are calculated based on the actual and predicted stock prices and table 2 illustrates the summary of accuracy test results.

Accuracy Percentage = 100 - [(|Actual Price - Predicted Price| / Actual Price) * 100] (1)

TABLE II. ACCURACY TEST RESULTS SUMMARY

Stock Symbol	Model	Average Accuracy
AAPL	ARIMA	98.04%
	LSTM	97.83%
	RL Agent	94.28%
	Ensemble	97.46%
MSFT	ARIMA	98.15%
	LSTM	94.65%
	RL Agent	91.39%
	Ensemble	96.74%
ORCL	ARIMA	97.89%
	LSTM	97.53%
	RL Agent	73.65%
	Ensemble	97.17%

According to the above preliminary results of this study, it indicates that the ARIMA model outperforms the other models in the short test run of one week, however when it comes to long-term prediction, ARIMA model struggles to identify the intricate patterns and make the predictions compared to the LSTM model. RL Agent model performance is not at an acceptable level since the limitations had when training the model such as resource limitations, training data limitations etc. However, as a combination of these separate models, ensemble model as an average shows more than 97% accuracy. While this model may not serve as a definitive investment solution, it stands as a testament to the potential of machine learning and data analysis in improving our understanding of financial markets.

VI. FUTURE WORK

To enhance this model to make better predictions, we can consider some key future improvements. First, we should aim to make the model to train daily to give predictions in real-time where the current model has to be trained manually after collecting daily data from yfinance. This training process has to be done manually at this point mainly due to resource limitations and the time it takes to train the models. But if we can make it learn from the most recent data every day, it could be more accurate and helpful. Next, we can expand the range of stocks it can predict. Currently, it only looks at Apple (AAPL), Microsoft (MSFT), and Oracle (ORCL). It would be great if users could choose any stock they want to predict. This would make the model more versatile. Getting more data is also important. We should try to find more historical news data related to each and every stock where currently it only considers 1.5 years of data which is not enough. We can also try different things like using more technical indicators and looking at what people are saying on social media, like Twitter, Threads etc. These additions could improve the model's accuracy. Lastly, it was possible to more powerful computers, it would make the whole process faster and more efficient. This means we could try out more ideas and make the model even smarter. These changes would make the stock price prediction system more useful and reliable.

VII. CONCLUSION

In conclusion, this research presents a comprehensive ensemble approach for stock price prediction, combining machine learning,

reinforcement learning, and NLP techniques. Overcoming challenges faced throughout the journey, such as model dependency on training data, underscored the importance of an adaptive approach in machine learning model development. The future improvements outlined some key aspects by setting a roadmap for enhancing system versatility. The real-world application acknowledges the inherent uncertainties in stock market investments by providing a tool for informed decision-making rather than a definitive solution. With a keen awareness of the project's limitations and potential future advancements, this research project stands as a testament to the synergy of machine learning, reinforcement learning, and sentiment.

ACKNOWLEDGMENT

I would like to express my heartfelt gratitude to all the individuals who supported me in my research journey, especially my supervisor Dr. Rasika Ranaweera for the unwavering support, guidance and motivation that he gave me throughout this journey. Also, I would like to thank Mr. Sumal Sugathadasa who is a Senior Software Architect as LSEG Technology, for the invaluable support, guidance, encouragement that he gave me throughout this research. Last but not least, I would like to thank my friends, family and lecturers for their unwavering support throughout my research journey.

REFERENCES

[1] Chong, E., Han, C. and Park, F.C. (2017) 'Deep learning networks for stock market analysis and prediction: Methodology, data representations, and case studies', Expert Systems with Applications, 83, pp. 187–205. Available at: https://doi.org/10.1016/j.eswa.2017.04.030

[2] Sezer, O.B., Gudelek, M.U. and Ozbayoglu, A.M. (2019) 'Financial Time Series Forecasting with Deep Learning : A Systematic Literature Review: 2005-2019'. Available at: http://arxiv.org/abs/1911.13288.

[3] Hegazy O, Soliman OS, Salam MA (2014) A machine learning model for stock market prediction. Int J Comput Sci Telecommun 4(12):16–23

[4] Shen S, Jiang H, Zhang T (2012) Stock market forecasting using machine learning algorithms. Department of Electrical Engineering, Stanford University, Stanford, pp 1–5.

[5] Chen L, Qiao Z, Wang M, Wang C, Du R, Stanley HE (2018) Which artifcial intelligence algorithm better predicts the Chinese stock market? IEEE Access 6:48625–48633.

[6] Yetis Y, Kaplan H, Jamshidi M (2014) Stock market prediction by using artifcial neural network. In: IEEE WAC, pp 718–722

[7] Ou P, Wang H (2009) Prediction of stock market index movement by ten data mining techniques. Mod Appl Sci 3(12):28

[8] Urolagin S (2017) Text mining of tweet for sentiment classifcation and association with stock prices. In: IEEE ICCA, pp 384–388

[9] Chakraborty P, Pria US, Rony M, Majumdar MA (2017) Predicting stock movement using sentiment analysis of Twitter feed. In: IEEE 6th international conference ICIEV-ISCMHT, pp 1–6

[10] Khatri SK, Srivastava A (2016) Using sentimental analysis in prediction of stock market investment. In: IEEE 5th international conference ICRITO, pp 566–569

[11] Yan D, Zhou G, Zhao X, Tian Y, Yang F (2016) Predicting stock using microblog moods. J China Commun 13(8):244–257

[12] Zhou Z, Zhao J, Xu K (2016) Can online emotions predict the stock market in China? In: international conference on web information systems engineering, pp 328–342

[13] Dang LM, Sadeghi-Niaraki A, Huynh HD, Min K, Moon H (2018) Deep learning approach for short-term stock trends prediction based on two-stream gated recurrent unit network. IEEE Access 6:55392–55404

[14] Vargas MR, dos Anjos CEM, Bichara GLG, Evsukof AG (2018) Deep learning for stock market prediction using technical indicators and fnancial news articles. In: IEEE international joint conference IJCNN, pp 1–8

[15] Chen W, Yeo CK, Lau CT, Lee BS (2017a) A study on real-time lowquality content detection on Twitter from the users' perspective. PLoS ONE 12(8):e0182487

[16] Li Q, Wang T, Li P, Liu L, Gong Q, Chen Y (2014a) The efect of news and public mood on stock movements. J Inf Sci 278:826–840. https://doi.org/10.1016/j.ins.2014.03.096

[17] Jeon S, Hong B, Chang V (2018) Pattern graph tracking-based stock price prediction using big data. J Future Gener Comput Syst. https ://doi.org/10.1016/j.future.2017.02.010

[18] Vijh, M. et al. (2020) 'Stock Closing Price Prediction using Machine Learning Techniques', in Procedia Computer Science. Elsevier B.V., pp. 599–606. Available at: https://doi.org/10.1016/j.procs.2020.03.326

[19] Khan, W. et al. (2022a) 'Stock market prediction using machine learning classifiers and social media, news', Journal of Ambient Intelligence and Humanized Computing, 13(7), pp. 3433–3456. Available at: https://doi.org/10.1007/s12652-020-01839-w

[20] Nikou, M., Mansourfar, G. and Bagherzadeh, J. (2019) 'Stock price prediction using DEEP learning algorithm and its comparison with machine learning algorithms', Intelligent Systems in Accounting, Finance and Management, 26(4), pp. 164–174. Available at: https://doi.org/10.1002/isaf.1459

[21] Khan, W. et al. (2022b) 'Stock market prediction using machine learning classifiers and social media, news', Journal of Ambient Intelligence and Humanized Computing, 13(7), pp. 3433–3456. Available at: https://doi.org/10.1007/s12652-020-01839-w

[22] Mehta, P., Pandya, S. and Kotecha, K. (2021) 'Harvesting social media sentiment analysis to enhance stock market prediction using deep learning', PeerJ Computer Science, 7, pp. 1–21. Available at: https://doi.org/10.7717/peerj-cs.476

[23] Kordonis, J., Symeonidis, S. and Arampatzis, A. (2016) 'Stock price forecasting via sentiment analysis on Twitter', in ACM International Conference Proceeding Series. Association for Computing Machinery. Available at: https://doi.org/10.1145/3003733.3003787.

[24] Mehtab, S., Sen, J. and Dutta, A. (no date) Stock Price Prediction Using Machine Learning and LSTM-Based Deep Learning Models

[25] Porshnev, A., Redkin, I. and Shevchenko, A. (2013) 'Machine learning in prediction of stock market indicators based on historical data and data from twitter sentiment analysis', in Proceedings - IEEE 13th International Conference on Data Mining Workshops, ICDMW 2013. IEEE Computer Society, pp. 440–444. Available at: https://doi.org/10.1109/ICDMW.2013.111.

[26] Li, Y., Ni, P. and Chang, V. (no date) Application of Deep Reinforcement Learning in Stock Trading Strategies and Stock Forecasting.

[27] Wadi, S. AL, Almasarweh, M. and Alsaraireh, A.A. (2018) 'Predicting Closed Price Time Series Data Using ARIMA Model', Modern Applied Science, 12(11), p. 181. Available at: https://doi.org/10.5539/mas.v12n11p181.

[28] Alwadi, R. (2015). Forecasting short term financial data. European Scientific Journal, 11(25), 251-255.

[29] Pai.P.F., & Lin.C.S.(2005). "A hybrid ARIMA and support vector machines model available in stock price forecasting". Omega,33(6), 497-505.

[30] Babu. C. N., & Reddy. B. E. (2012, March). "Predictive data processing on average global temperature using variants of ARIMA models." In IEEE-International Conference On Advances In Engineering, Science And Management (ICAESM-2012) (pp. 256- 260). IEEE

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 is a significant issue faced by Tissue culture laboratories in Sri Lanka, resulting in considerable plantlet wastage and production loss. This research paper aims to automate the identification of fungi and bacteria contaminants that appear inside banana cultures. The primary objective is to develop an innovative system leveraging automation, machine learning, and image processing technologies for the precise identification of these contaminants. The system's design emphasizes user-friendly interfaces, enabling non-technical individuals to identify unknown fungi and bacteria contaminants with minimal assistance. By streamlining the identification process, the research project intends to provide laboratories with a practical tool for efficient and easy recognition of contaminants. By integrating machine learning algorithms and image processing, the proposed system offers a revolutionary approach to contamination identification.

Keywords*—Machine Learning, Image Processing, Transfer Learning, Fungi, Bacteria, Tissue culture, Banana, Banana cultures, Microscopic Images, 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 rapid multiplication of uniform plantlets out of a single mother plant [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 faced in the tissue culture industry locally as well as 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 never visible at the initial stage of a culture, but some internal contaminants emerge in later subculture phases or after 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 [6][7].

Some laboratories in Sri Lanka remove all the cultures affected with contamination from laboratories as there's the possibility of other healthy plants getting affected by the same fungi or bacteria even though the plants are residing inside tightly closed jars. Specially fungi contain microscopic spores, and they can transmit through air and affect other plants. Even if the plants are inside tightly closed sterilized jars, there is a possibility of gaps in the lid that allow gaseous exchange with the outside environment [8]. Also, there's a possibility that cultures get affected by the operators who handle these plants as the nutrient media that is used to grow the banana plantlets contains lots of sugar, is very sensitive and can be a breeding ground for many fungi and bacteria [9]. Since they lack a feasible methodology to identify these contaminants, they do have an issue in deciding what sort of antibiotic or fungicide to precisely use on that particular fungi or bacteria to mitigate them and further spread among the cultures. Due to the lack of a less costly reliable identification procedure, laboratories face lots of production losses yearly [1].

There is much research done 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 which contains different kinds of species such as fungi, bacteria, protozoa, yeast, viruses, etc. Therefore, no system has been developed, or research has been conducted to capture or identify the whole spectrum of these species. Some research papers have mentioned systems designed to specifically identify different strains of the same species, or for specific groups (around 5 or 6 species) of microorganisms [11]. However, there has been no system developed to automate the identification of fungi bacteria species in banana tissue culture plants. By having this kind of system in a laboratory, it will be feasible for them to decide on what are the tailored specific mitigation steps that they will take to improve their 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 intends to provide a feasible solution via a system based on machine learning and image processing approaches to identify species of fungi and bacteria contaminants related to banana tissue culture plants.

II. METHOD

A. *Materials*

Require a microscope that has a minimum of 400x magnification and an external camera or a smartphone camera with a clear resolution that could be mounted on top of the microscope to capture the microscopic images of microorganisms. By placing a sample in a slide, observe and adjust the microscope until receive a clear image. 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" function.

The captured image will be uploaded to the system via the front end. The system will then process the image and generate an output stating to which species the input microorganism belongs. A detailed scientific (taxonomy) classification is provided along with fungicide/antibiotic recommendations that can apply to the identified fungi or bacteria.

B. *Dataset*

In order to train the classification model, fungi and bacteria images from the DIBaS (Digital Image of Bacterial Species)[13] dataset and publicly available microscopic images from various surveys, articles, documents, etc. 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, that is, from the total dataset, 80% is used for training while 20% is used for validation.

C. Image Processing

Image augmentation is applied to make the training dataset more balanced as the current dataset is imbalanced. In order to achieve this, various image augmentation techniques such as rotation, flipping, etc, are applied to create several variations of the existing images. We have used the Python library 'ImageDataGenerator' from Keras to perform data augmentation on the dataset.

Images belonging to *penicillium* have been augmented by augmentation factor = 2, *proteus* by 5 and *staphylococcus* by only 1. The difference between the augmentation factor for different classes is due to the reason of making the dataset balanced. Also, all the training images have been normalized for their pixel values to be in the [0,1] range to make it easier to process the images. The same normalization technique has been applied to the validation image data as well.

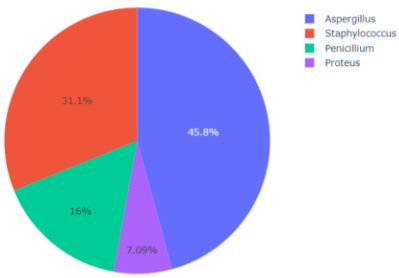


Figure 3. Original dataset class distribution related to the 4 classes: Aspergillus, Penicillium, Proteus and Staphylococcus.

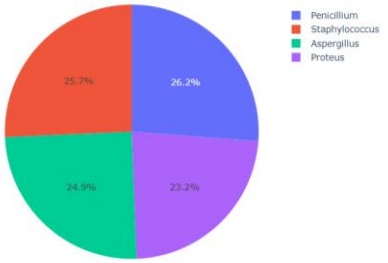


Figure 4. Class distribution after augmenting the dataset.

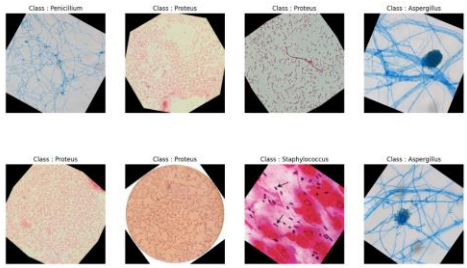


Figure 5. Visualizing sample images out of the augmented image dataset with class labels

D. Image Classification Model

In order to create the bacteria-fungi image classification model, a set of technologies and techniques such as Image augmentation, data visualization, and Transfer Learning were used. Transfer learning is a machine learning technique where we customize and use pre-trained models developed for a particular 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 smaller 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, InceptionV3 is used. InceptionV3 is a deep convolutional neural network (DCNN) architecture that is mostly

used for transfer learning [7], [10], and it is a pre-trained model on a large dataset called ImageNet [14]. In order to train the model, we used the Kaggle Notebooks along with their available GPU accelerators. A Global Average Pooling (GAP) layer, a fully connected layer with ReLU activation and a final dense layer with softmax activation for classification are added to the architecture. The Adam optimizer is used to assemble the model with a specified learning rate. To save the best weights during training, two callbacks are used: EarlyStopping (patience of three epochs) and ModelCheckpoint. The model is then trained for 15 epochs on the provided datasets (photos with dimensions of 500x500), with the training history saved in the 'history' variable for further analysis.

E. System Development

A web application was developed according to the MVC (Model-View-Controller) architecture by integrating the trained classification model. The frontend was developed using React JS and the backend using Python with Flask framework. Flask has been used to receive HTTP requests of images from the frontend and reciprocate JSON responses back to the frontend. A log file has been set up to record all the activities that are happening in the application, so it is easier to investigate if an error occurs.

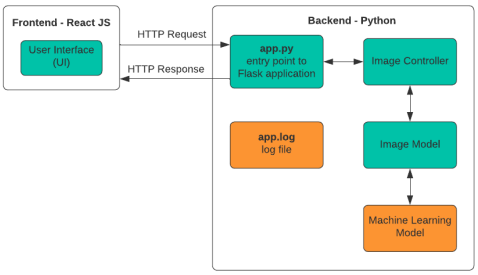


Figure 6. System workflow of the developed web application

Image upload functionality is provided to users to browse their device files and upload an image. To ensure accurate predictions, several validations have been set such as, the Image quality should be greater than 500x500 pixels and the size must be less than 20MB. Also, only JPG/PNG image file formats are supported. Additionally, an integrated camera functionality was provided so users have the option to capture images from the application itself.

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 having a smaller dataset, the reason for achieving a comparatively high accuracy is due to the use of image augmentation techniques and transfer learning where a pre-trained model (Inception V3) was used.

Using the saved data in the history variable which was used while training the model, the learning curve for the trained model is plotted. As represented in Figure 5, the training loss has decreased in each epoch is a good sign, indicating that the model is learning and improving its predictions on the training data. The validation loss fluctuating and sometimes increasing can indicate that the model might be overfitting. The reason for overfitting mostly occurred due to the comparatively smaller size of the dataset. Acquiring more diverse images for training may help overcome overfitting. The classification model becomes too specialized in fitting the training data and does not generalize well to unseen data (validation data in this case).

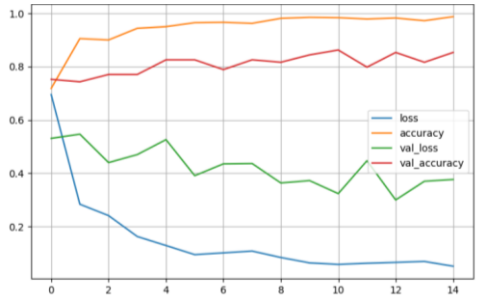


Figure 7. Plotted learning curve for the trained classification model.

The final developed system was tested for errors and its performance. Notably, the system responded to user inputs well. The upload page functionalities, including image selection, upload, and camera use, performed as expected. The successful display of the input image, accurate classification results, and associated remedies on the results page validated the APIs are performing as expected by retrieving essential data from the database and sending it to the frontend. As well as it showed that 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 modern need for adaptability across various devices.

IV. Future work

Future recommendations include both improving existing features and expanding its capabilities. As of now, the model has been trained focusing on a single crop, i.e., 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*, Orchids, etc., increasing the system's adaptability and applicability in a variety of laboratory settings. In future work, it is possible to overcome the model overfitting by acquiring more diverse images for training for a single species.

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 plants, 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 emphasizing user-friendly interfaces. Despite obstacles such as dataset limitations and imbalanced data, approaches such as transfer learning and image augmentation were used. This study is a significant step toward revolutionizing microbial identification in tissue culture, offering enhanced efficiency and resource conservation for laboratories.

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, Mr. Pramudya Thilakarathne, 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 in Sri Lanka 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.

References

[1] 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. [Online]. Available: http://www.ijser.org

[9] "Different types of contaminants in plant tissue culture - Lab Associates.". [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," Archives of Computational Methods in Engineering, vol. 29, no. 3. Springer Science and Business Media B.V., pp. 1801–1837, May 01, 2022. 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." [Online]. Available: https://meta-album.github.io/datasets/BCT.html

[14] "ImageNet." [Online]. Available: https://www.image-net.org/

A Smart Blood Inventory Management Platform for the National Blood Transfusion Service of Sri Lanka with Blood Demand Prediction using Machine Learning

Sanuli Weerasinghe
Faculty of Computing
NSBM Green University
Pitipana, Homagama, Sri Lanka
sbweerasinghe@students.nsbm.ac.lk

Rasika Ranaweera
Faculty of Postgraduate Studies & Professional Advancement
NSBM Green University
Pitipana, Homagama, Sri Lanka
ranaweera.r@nsbm.ac.lk

Dulanjali Wijesekara
Faculty of Computing
NSBM Green University
Pitipana, Homagama, Sri Lanka
dulanjali.w@nsbm.ac.lk

Abstract* – Effective blood inventory management is crucial for ensuring a secure and efficient supply of blood and blood products within the healthcare system. Managing blood inventory however, poses a set of unique challenges, necessitating a distinctive approach tailored to the intricacies of the blood supply chain. In response to these complexities, this research proposes a Smart Blood Inventory Management Platform, employing a predictive approach designed specifically for the National Blood Transfusion Service of Sri Lanka to address the complex dynamics of Blood Demand and Supply Chain. By leveraging predictive analytics, the platform aims to mitigate issues related to overstocking and understocking, enhancing the overall efficiency of blood inventory management. Underlining the study, a prediction model is presented for the monthly blood demand prediction of the 8 blood types, including an evaluation of its performance using metrics such as Root Mean Square Error (RMSE), Mean Absolute Percentage Error (MAPE), and Coefficient of Determination (R^2). The results indicate promising outcomes, with the predictive model demonstrating a trend of improved accuracy in forecasting blood demand. Hence, this research will contribute to the domain of blood supply chain management by emphasizing the role of data-driven decision-making and predictive modeling.

Keywords* - Blood Supply Chain, National Blood Transfusion Service, Blood Inventory Management, Overstocking, Understocking, Demand Prediction, Machine Learning, Decision Tree Regression.

I. INTRODUCTION

Blood is a vital fluid that plays a fundamental role in sustaining life at numerous crucial circumstances, particularly in moments where the natural supply of blood within the body is insufficient or compromised. The productiveness of a healthcare system depends significantly on the availability of a safe and ample blood supply. The National Blood Transfusion Service (NBTS) holds the central position in this delicate balance, by ensuring that a sufficient blood supply is available to meet the national demand without any crises.

The Blood Demand & Supply Chain is the process and network involved in managing the supply and distribution of blood and blood products to meet the demand. This chain can be conceptualized as a complex system consisting of four interconnected echelons: Collection, Production, Inventory and Distribution [2]. The supply chain begins with voluntary blood donations. Collected blood undergoes processing into whole blood and components. Inventory management involves maintaining and controlling suitable

stock levels and monitoring expiration. Distribution ensures timely allocation of blood and blood products to the hospital network.

Despite the role of the NBTS, challenges persist in realizing an optimized Blood Demand & Supply Chain. The equilibrium is often disrupted by factors such as the perishable nature of blood, criticality of supply, stochastic behavior of blood demand and supply, and diversity of available blood types, leading to instances of overstocking or understocking. These inefficiencies in blood inventory management can have severe consequences in the quality of patient care. Hence, addressing these challenges is crucial for the entire healthcare landscape, emphasizing the need for a robust and intelligent solution.

Considering the unique complexities in blood inventory management, typical approaches such as Just-in-Time (JIT) principles aren't feasible, mainly due to the perishability of blood and the complex logistical processes involved, including collection, testing, processing and distribution, which require sufficient lead time. Thus, the necessity to investigate on a tailored inventory management approach which can specifically address the unique challenges presented in the context of blood inventory is identified.

Proper planning of demand is the key to achieve good inventory management. Demand Planning is the process of predicting future demand and translating those predictions into actionable steps aimed at fulfilling the identified demand [8]. The foundation for supply chain planning is accurate demand prediction [9] and therefore, predictive approach emerges as the key strategy for blood inventory management. The proposed system equips the NBTS with analytics on monthly blood demand prediction for each blood type which will result in optimization of the inventory levels, by reducing the risk of shortages and ensuring timely responses to critical demands.

II. Literature Review

In this section, existing systems will be reviewed which are been proposed in the field of blood inventory management and blood supply chain optimization.

A. Blood Bank Information System Based on Cloud Computing in Indonesia (2019)

The Blood Bank Information System (BBIS) is developed as an application that integrates hospitals, UTD PMI (Central Blood Transfusion Unit - Indonesian Red Cross), and individual donors across Indonesia. The BBIS encompasses various functions, including enabling hospitals to monitor the blood supply of UTD PMI and other hospitals, facilitating blood supply requests and direct communication among hospitals, UTD PMI, and individual donors, sending notifications to donors regarding blood donation events, recording all data in a secure cloud storage accessible to hospitals and UTD PMI (including blood supply data, medical history, donor history and donors' personal information) and generating push notifications for donor requests [7].

B. Blood Bank Management System in Malaysia (2015)

Blood Bank Management System is developed to manage blood bank of Hospital Sultanah Nur Zahirah Hospital (HSNZ). The purpose is to handle the stocks of the blood bank systematically by providing efficient management of blood bags received from the blood donation events. Through this system, results of the blood bags will be indicated of whether acceptable for transfusions or not and will also generate reports such as blood stock report, donor's gender report and the total blood donations per particular month/year [10].

C. Online Blood Bank Management System in Pakistan (2019)

The purpose of this web based Blood Bank Management System is to digitalize the blood bank operations and to facilitate the co-ordination between blood supply and demand. This system manages blood stocks and maintains information of donors, blood requests and the blood availability [6].

D. E-Blood Bank Application for Blood Transfusion Unit (2018)

This is a web-based application which serves the purpose of providing real-time blood stock availability information and to efficiently connect the Blood Transfusion Unit (BTU), recipients and donors. E-Blood Bank application encompasses various service features, including the processing of donor data, maintaining donation history, managing the blood donation data, handling blood requests and providing real-time stock information [5].

The research gap lies in the absence of a comprehensive solution that combines inventory model with demand forecasting capabilities. Therefore, this research aims to bridge this gap by providing a proactive and data-driven solution by utilizing machine learning algorithms to analyze historical blood demand patterns and identify significant factors influencing demand fluctuations. With the integration of predictive capabilities into this system proposed, it is expected to ultimately streamline the Blood Demand & Supply Chain.

III. Methodology

The research methodology accommodated in this project is Design Science Research Methodology, which is an approach to develop innovative solutions to real-world problems. It is a systematic procedure of creating practical solutions for solving problems, assessing the effectiveness of these solutions and communicating the results [1].

In the process of fact collection, qualitative methods particularly interviews have been used as the primary mechanism for capturing the domain specific knowledge. Eengaging in interviews with stakeholders at the NBTS representing different departments, enabled to gain a holistic understanding into the different aspects of the Blood Demand and Supply Chain which have been utilized in developing the proposed system.

For the development of the Smart Blood Inventory Management Platform, the SCRUM methodology is used, which is a widely adopted agile methodology that emphasizes iterative and incremental development wherein, the project is broken down in to manageable sprints.

IV. System Architecture

The proposed system is developed as a web-based system with a centralized database to store and organize all related data including donors, donations, requests, processed blood bags, donation events and demand predictions for each blood type. The system architecture is displayed in Fig 1.

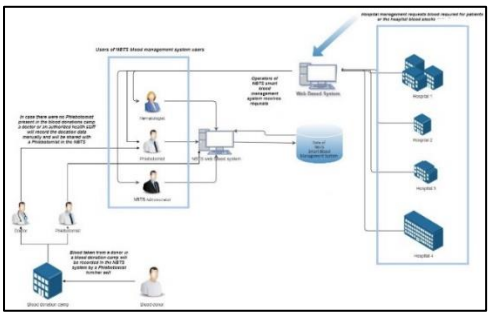


Fig 1. Proposed System architecture

The NBTS staff including the Admins, Phlebotomists, Haematologists, cluster centers and hospital staff will all be connected to the web application.

The Monthly Blood Demand Prediction is the most significant feature of the proposed system. It employs advanced data analysis and machine learning techniques to forecast the demand for whole blood at a monthly level. Fig 2 below depicts the workflow of the prediction model.

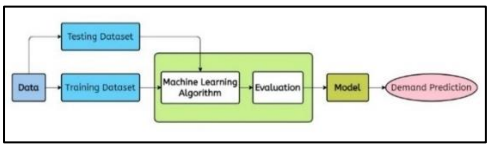


Fig 2. Workflow of the Prediction Model

V. Implementation

The Smart Blood Inventory Management Platform has been developed as a web application using PHP, Python, and MySQL.

By entering the credentials for the specific user account in the Home page, users

are navigated to the respective dashboard based on their user role.

The main functionalities of the system are donor registration, entry of donations, handling and tracking of processed blood bags, handling of blood requests, handling donation events, creation of user accounts and data-driven decision making through predictive analytics. Moreover, hospitals can place their requests for blood.

The machine learning algorithm chosen for the demand prediction task is Decision Tree Regression which is a model to forecast the value of a specific target variable based on various input factors [3]. This supervised learning approach is trained using a labeled dataset, wherein both the input variables and their corresponding output values (the variable to be predicted) are utilized to train the model.

The prediction dashboard is the most significant feature of the proposed system which is displayed in Fig 3 below.

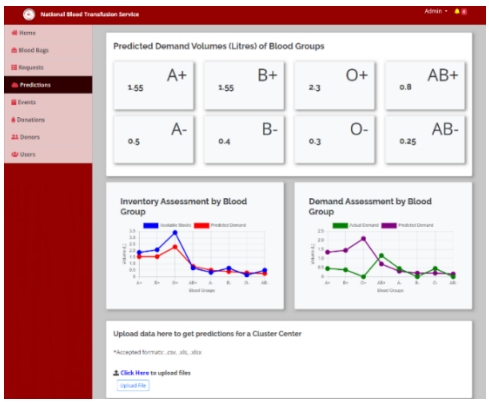


Fig 3. User Interface for Prediction Dashboard

Currently, the local demand prediction is considered for a particular cluster center with the expectation to expand to the national demand in future. Therefore, a file upload section is included, where the user is able to upload datasets for which they can get predictions for any other cluster center affiliated to the NBTS.

VI. RESULTS

To evaluate the performance of the prediction model and to assess the effectiveness of its forecasts, a performance analysis is conducted involving the computation of 3 metrics: Root Mean Square Error (RMSE), Mean Absolute Percentage Error (MAPE) and the Coefficient of Determination (R^2).

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(T_i - O_i)^2}{n}} \quad (1)$$

$$MAPE\ (\%) = \left(\frac{1}{n}\sum_{i=1}^{n}\frac{|T_i - O_i|}{T_i}\right) \times 100 \quad (2)$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(O_i - T_i)^2}{\sum_{i=1}^{n}(T_i - \bar{T}_i)^2} \quad (3)$$

Where O_i, T_i and $\bar{T}_i$ represents the predicted values, actual values and the Mean of the actual values respectively.

To assess the performance of this predictive model, the Coefficient of Determination (R^2) is considered, baselining that the model's predictive accuracy improves when its R^2 value approaches closer to 1 [4].

The Table I provides a summary of values obtained for above performance metrics after executing the prediction model for the 8 blood types.

TABLE III. STATISTICAL INDICATORS OF PERFORMANCE

Blood Type	RMSE	MAPE	R^2
O+	434.623	10.035	0.794
A+	796.111	76.869	0.641
B+	386.677	20.332	0.538
AB+	201.128	27.583	0.718
O-	89.652	11.460	0.744
A-	130.317	42.467	0.725
B-	31.544	4.346	0.942
AB-	31.741	10.332	0.935

Fig 4 below elaborates the results after executing the blood demand prediction model. This visualization showcases a

comparison between the actual demand and the predictions made by this model, represented individually for each blood type.

The results demonstrate promising outcomes, particularly evident in the trend of the R^2 values approaching closer to 1 across all blood types. The blood types O+, O-, A-, B-, and AB- are notably close to 1, suggesting a high degree of accuracy and reliability, whereas blood types A+ and O+ show room for improvement. This discrepancy may be due to the limited number of available records, which has likely influenced the model's precision.

The variability in R^2 values observed across different blood types can be attributed to several factors. Positive blood types being more common and prevalent, are likely to exhibit more fluctuating demand patterns, leading to comparatively lower R^2 values, as the predictive model may struggle to accurately capture these variations. Negative blood types being rarer, may have more stable and predictable demand patterns, resulting in higher R^2 values.

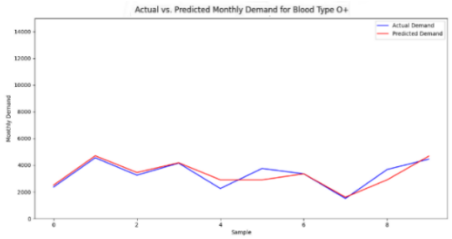
In the future, with the accumulation of additional historical data, the predictive capabilities are expected to enhance significantly, thereby providing the model with a broader dataset, enabling it to better capture the diverse variations in demand patterns across various blood types. This will assist NBTS to develop more targeted strategies for blood inventory management, with a focus on optimizing inventory levels for different blood types.

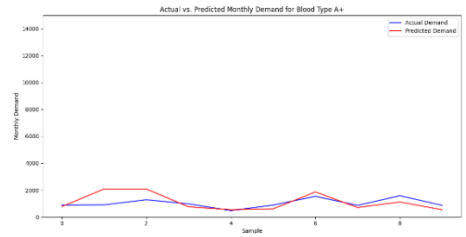
VII. Conclusion

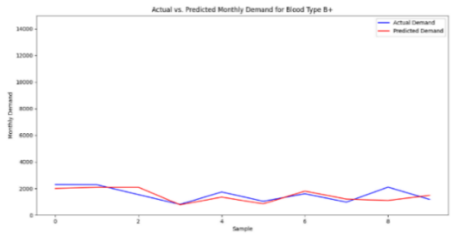
The primary aim of this study was to empower blood inventory management by integrating advanced technologies and predictive analytics into the existing processes of NBTS. The focus was to develop a comprehensive platform that ensures the efficient allocation of blood resources, minimizes wastage and ultimately enhances the quality of patient care. By leveraging machine learning and real-time data, this system predicts blood demand, reducing the likelihood of overstocking and understocking. The further enhancements suggested, will further elevate the platform and its impact on the healthcare landscape.

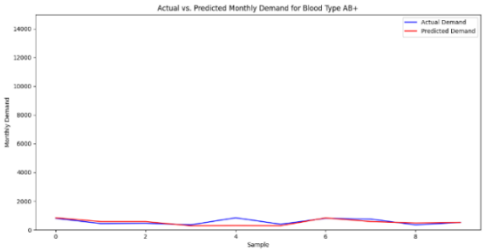
VIII. Future Works

Potential enhancements and features that can further elevate the functionality and efficiency of this Smart Blood Inventory Management platform are explored with the intention to make this system more of an ecosystem.









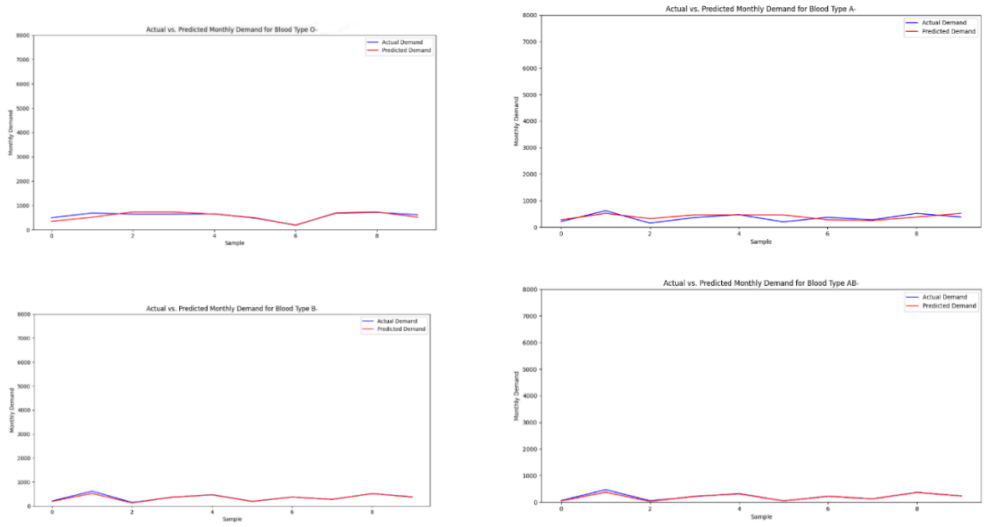


Fig 4. Actual Demand Vs Predicted Demand

- Implementing the system for the National Blood Transfusion Service of Sri Lanka.
- Automated data extraction from ID documents (NICs, licenses, donation cards) via image processing for donor registration.
- Handling urgent same-day orders for critical blood requests with priority.
- Tracking the status by hospitals for their routine blood requests.
- Real-time alert system for hospitals to receive status updates on their requests.
- Mobile app for donors: appointment scheduling, notifications on upcoming donation events and educational content.
- Integration of EHRs and hospital databases to predict blood demand based on patient admissions, surgeries and ongoing medical treatments.
- Advanced algorithms for national level blood demand prediction.

ACKNOWLEDGEMENT

The authors gratefully acknowledge the stakeholders at the National Blood Transfusion Service, Narahenpita, Sri Lanka for their invaluable support in deepening the understanding on the domain and providing the necessary datasets which greatly contributed to the success of this project.

REFERENCES

[1] Dresch, D.P. Lacerda, and J.A.V. Antunes, "Design science research," Springer International Publishing, pp. 67–102, 2015.

[2] A.F. Osorio, S.C. Brailsford, and H.K. Smith, "A structured review of quantitative models in the blood supply chain: a taxonomic framework for decision making," International Journal of Production Research, vol. 53, no. 24, pp. 7191–7212, February 2015. DOI = https://doi.org/10.1080/00207543.2015.1005766.

[3] Gupta, A. Rawat, Jain, A. Arora, and N. Dhami, "Analysis of various decision tree algorithms for classification in data mining," International Journal of Computer Applications, vol.163, no. 8, pp. 15-19, 2017.

[4] Chicco, M.J. Warrens, and G. Jurman, "The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation," PeerJ Computer Science, vol. 7, pp. 623, 2021.

[5] L. Sumaryanti, S. Suwarjono, and L. Lamalewa, "E-Blood Bank Application For Organizing and Ordering The Blood Donation," International Conference on Science and Technology (ICST 2018) Atlantis Press, vol. 1, pp. 754–758, 2018.

[6] M. Hamza, J. Ghaffar, G. Ibrahim, and M.I. Ullah, "On Line Blood Bank Management System: A Web Application," Journal of Information Engineering and Applications, vol. 9, no. 6, October 2019.

[7] M.N.S. Ramadhan, A. Amyus, A.J. Fajar, S. Sfenrianto, A.F. Kanz, and M.S Mufaqih, "Blood Bank Information System based on cloud computing in Indonesia," Journal of Physics: Conference Series, 2019. DOI = https://doi.org/10.1088/1742-6596/1179/1/012028.

[8] S. Hamadneh, O. Pedersen, M. Alshurideh, B. Al Kurdi, and H.M. Alzoubi, "An Investigation of The Role of Supply Chain Visibility Into The Scottish Blood Supply Chain," Journal of Legal, Ethical and Regulatory Issues, vol. 24, September 2021.

[9] S.M. Fortsch and E.A. Khapalova, "Reducing uncertainty in demand for blood," Operations Research for Health Care, vol. 9, pp.16-28, 2016. DOI = http://dx.doi.org/10.1016/j.orhc.2016.02.002.

[10] S. Sulaiman, A.A.K.A. Hamid, and N.A.N Yusri, "Development of a Blood Bank Management System," Procedia - Social and Behavioral Sciences World Conference on Technology, Innovation and Entrepreneurship, pp. 2008–2013, 2015. DOI = https://doi.org/10.1016/j.sbspro.2015.06.215.

Community-Based Online Charity Platform

Nethmi Weththasinghe
Faculty of Computing
NSBM Green University
Colombo, Sri Lanka
nwetthasinha@gmail.com

Mohamed Sapraz
Faculty of Computing
NSBM Green University
Colombo, Sri Lanka
shafrz@nsbm.ac.lk

Abstract*— The research aims to design and develop a user-centric platform that improves user experience, ensures accessibility and security, improves transparency, integrates social sharing tools, donates through various payment methods and volunteer for any project. There are many people who need social support in different ways. It can be informational, emotional or practical support. But current charity organizations rely on one linear goal and they are not focusing all areas. It is handled and managed by individuals, organization or government. Also, lack of transparency is one of the limitations of current organizations. Many charity organizations do not use technology in their work. There are several charity organizations that use technology for manage charity activities as well. Here, through questionnaire-based data collection and a rigorous review of findings, the purpose of this research is to provide practical insights and recommendations for platform developers, academics and stakeholders involved in online charity communities. These findings provide actionable insights for improving existing platforms and motivate further exploration in this dynamic and evolving field.

Keywords—Social support, Charity community, Online charity, Donate, Volunteer

I. INTRODUCTION

The various forms of assistance and resources that individuals receive from their social networks are referred to as social support. This support can come in many different forms such as emotional support, instrumental support, companionship support, informational support and esteem support [1]. Emotional support include sympathy, relationships, encouragement, empathy. Instrumental support means providing tangible assistance. Informational support means giving advices, sharing information. Esteem support include promote someone's skills and ability [2]. Social support can also be categorized by age, gender, health and more facts.

Due to ongoing economic crisis in Sri Lanka, the under privileged people are suffering a lot by this situation. There are many children who have stopped their education especially because of those conditions. And also, many schools in small villages do not have facilities for children to study. Therefore, charity organizations can organize seminars, workshops and events to motivate children [3]. Not only the economically but also many people are suffering mentally as well. Some people need counselling or advice. Some people need money, food or good. Some people need knowledge or education. So, different people need different things.

Nowadays, charitable giving is an important part of society in Sri Lanka, but it can be challenging to connect donors with the causes they care about. And also, it causes a disconnect between donors and beneficiaries because traditional charity models lack personal engagement and community involvement. Many charity organizations don't use new technologies for it. And very least number of organizations use new technologies such as using web or mobile applications, money transferring, card payment to serve. The current charity organizations based on one linear goal. And

also, in traditional organization rely on individuals, organization or government. Lack of transparency is another limitation of traditional charity organizations. Many organizations do not provide a proper plan for charities which leads to a breakdown of trust between donors and the organization. There are so many scammers as well [4]. And also, most of organizations do not focus on informational support or emotional support. The purpose of this research study is to investigate the potential of community-based charity platforms to address the challenges of traditional charity models.

The study aims to identify the key features and functionalities of successful community-based charity platforms and how they can foster engagement and participation among donors and beneficiaries. The expected outcomes of the research include insights into the challenges of charitable giving, the impact of community engagement on sustainability and effectiveness and recommendations for measuring and tracking the success of community-based charity platforms.

II. Literature Review

The purpose of this literature review is to explore existing research on community-based charity platforms in Sri Lanka, focusing on their unique characteristics, impact on local communities and challenges they face.

A. *Progressive Learning Design Strategy to Improve Impact Maturity of Charity organizations [5]*

There are so many people who need social support in different way. And also, there are many small scale, informal and unregistered charity organizations. They focus on one goal and different organization have different characters and needs.

B. *Development of a Web-Based Charity Organizations and Donation Management System [6]*

According to Inland Revenue Department Website, there are 84 approved charities in Sri Lanka. Around 5 of them are only provide online donate facility. Few organizations provide donate offline via bank transferring. Around 23 organizations have websites but they only provide view it. And also, there are about 42 organizations that do not have websites or do not use technology. According to this, the number of online donation charities in Sri Lanka is low.

C. *The Strategic Use of Information Technology by Nonprofit Organizations: Increasing Capacity and Untapped Potential [7]*

Nonprofits organizations must improve their organizational capabilities in the areas of long-term IT planning, budgeting, staffing and training, performance measurement, internet and website capabilities and senior management vision, support and involvement to maximize the mission-related impact of Information Technology.

D. *Development of a crowdfunding web-platform to improve the effectiveness of medical fundraising for a charity fund [8]*

It has been a common development process divided into clients and server sides. This investigation begins by introducing the technologies employed in the models. Bootstrap, MaterialUI and Font-awesome are used to style the contribution site model. They may be used by providing their website's import link in their code. The user interface design is created using Figma web design tool. React with JavaScript is used for front end development framework. C# language have been used for server side and SQL server has been used for database.

E. *Platform for Tracking Donations of Charitable Foundations based on Blockchain Technology [9]*

Blockchain technology is currently being used across a wide range of sectors. This technology is a potential feature that comes to the system. Blockchain technology enables to make the contribution and funding transaction processes transparent. To track all information about donations, transactions and donors, a single platform for tracking donations must be established. This article examines the description of the platform for

monitoring donations based on blockchain technology. The system which is based on blockchain technology, allows transparent accounting of activities for donors, charitable foundations and recipients. The charitable platform should provide a transparent donation route, allowing public users and contributors to track and monitor where, when, and to whom charity donations were disbursed.

III. Methodology

Design Science Research Methodology (DSRM) is the systematic approach used for this research study. There are many charity organizations and communities in Sri Lanka today with different limitations. In this research focus on develop a community-based charity platform with preventing those limitations. Current charity organizations rely on a single linear goal and are managed by an individual, organization or government. People cannot do charity with what they have. And those traditional charity organizations have lack transparency. People can donate money but people don't know what happens inside and can't directly connect to charity. Basically, it could be improving donor engagement, enhancing transparency in donations or streamlining volunteer management. But as a community there is no single platform to help everyone in every way. The specialty of this system is that everyone can join the charity as a community and volunteer with their interest and ability. For that should collect data to verify it and get data from user experience and preference.

A. Fact collection mechanism

There are two types of research paradigm that can be used. Positivist research paradigm focus on quantitative method and interpretivist focus on qualitative method. In this research, it will use Positivist research paradigm to gather quantitative data through this questionnaire (Google form). Preparing questionnaires is an easy and effective way to collect data from a large audience. These data are highly accurate and applicable to the study area. In this research, 100 responses are collected for data analysis using questionnaire. Male and female, young and old, students, employees and unemployed all participated in this data collection process. Through this questionnaire, anyone can get a good idea about this research area.

B. Result and solution

According to the result got from the questionnaire,

- 94% have donated using technology and 6% have not donated using technology. Therefore, it is would be better to develop an online platform to do charity because of their busy life.
- The most popular technology to donate is money transfer. 85.1% donated through money transfer. 85.1% donated through cash transfer. 44.7% are donated through star points. 68.1% donated via card payment and 1.1% via crypto. Therefore, it would be better to add various payment methods to donate on this online platform.
- The charity organizations pay less attention to emotional support. Therefore, it is better to add different projects for these supports.
- Many charity organizations give money than other forms of support. Many charitable organizations provide food and physical health as a service as well. Fewer charitable organizations provide education, advice, jobs offering and mental health.
- 40% of the target population say they cannot connect with beneficiaries through current charities and 49% say they are not sure. Therefore, it is better to create an opportunity to improve the connection between volunteers and beneficiaries.
- 62% people say that charity organizations have lack transparency. 35% people are not sure about this and 3% people say that charity organizations have

more transparency. Therefore, details and clear records should be shown such as how funds are spent and how much is available to improve transparency.

- 61.8% people says web application is more suitable and 8.8% says mobile application is most suitable. Also, 29.4% says both will suit. Therefore, a web application is the most suitable than mobile application for charity-based online community.

After the study of the previous literature and data analysis part it is possible to get an idea that the predicted problem exists in the society and the target users need this charity based online community platform. According to this, decided the requirements that need for this platform. Functional requirements and non-functional requirements are as follows. Functional requirements are key features of the system and without them such a platform cannot be designed or developed. Non-functional requirements are quality attributes and having those qualities isn't mandatory, but it's good to have them.

Functional requirements

- Authenticate user when login or signup
- View profile and update password
- Become a volunteer
- View volunteers
- View projects
- Can donate through different payment methods
- View donors related to different projects
- View funds handling details

Non-functional requirements

- High security
- Scalability
- Good performance
- Reliability and availability
- Privacy and data protection
- Cost effectiveness
- User friendly

IV. Implementation

After determining the required features this charity based online community platform can be designed and developed. Some features can be accessed without logging in, such as viewing projects and donating. There are also some features that require logging in for access, such as donating via card payments, becoming a volunteer, and viewing volunteers. All users can log in with the same account. Any user can be a donor, volunteer or beneficiary. One user can be a volunteer for one project and the same user can be a beneficiary for another project. Moreover, the platform will truly help donors and volunteers to donate and volunteer with high transparency and get any kind of support to the beneficiaries. Most users need a web application for that and the users do not need to install any application. So, a web application should be designed and developed.

Here is the flowchart related this proposed system.

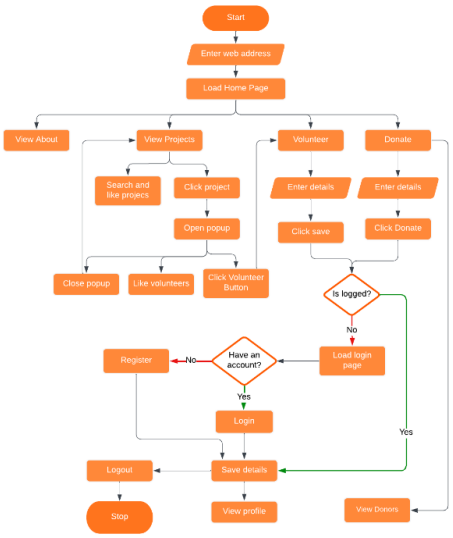


Fig. 1. System flow chart

When it's comes to the proposed system development, Angular, Spring Boot and MySQL was used. For the API testing, Postman and for the front-end testing, google chrome web browser was used.

As per requirements, these are some user interfaces designed and developed.

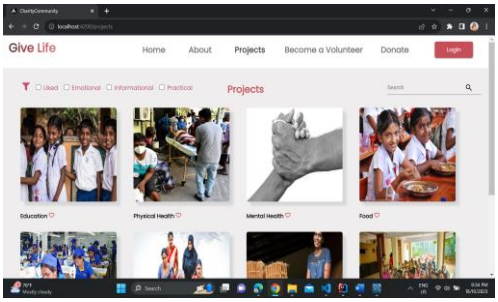


Fig. 2. View projects UI

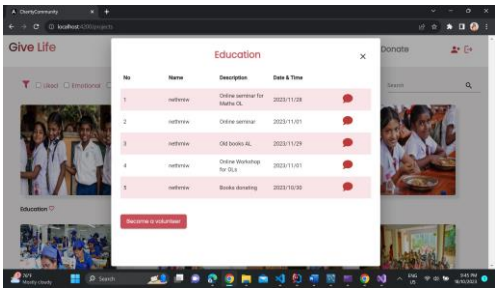


Fig. 3. View volunteers UI

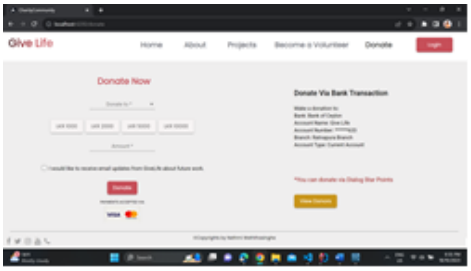


Fig. 5. Donate UI

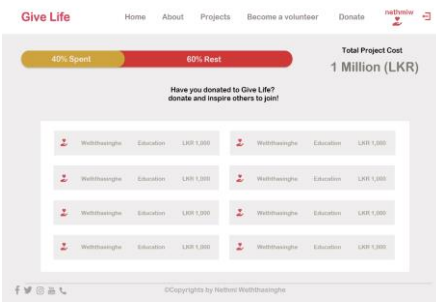
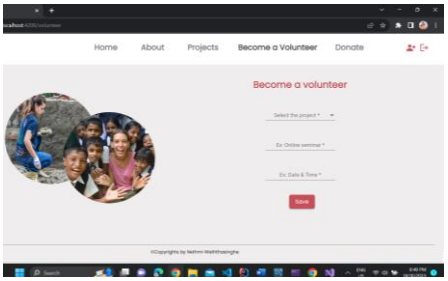


Fig. 6. Donors UI

V. DISCUSSION

Research findings and how they align with the originally defined objectives were broadly discussed. Each research objective and present key insights and findings related to it were revisited. One objective is to identify how many people are willing to engage with charity with technology. 94% of people of the target population use technology for charity and most of them use money transfer to donate. The research shows high levels of willingness, it underscores the potential for success in an ICT-supported charity platform with various payment methods.

Another objective is to identify what do people want as a help other than money. It seeks to uncover the various forms of support or contribution that individuals value beyond financial donations. This may include time, skills, resources or other forms of support. By



identifying and understanding these alternative modes of support, the research informs the development of an ICT-supported charity platform that can engage the charity with what they have.

Another objective is to design and develop an ICT-supported community-based charity platform that can connect directly with donors, volunteers and beneficiaries and engage the charity with what they have. The primary purpose of the platform is to create direct connections between donors, volunteers and beneficiaries and to give users the opportunity to connect with the charity using their unique assets and capabilities. Therefore, there should be more transparency in the design and development of the system

and should be able to connect everyone together as a community.

Final objective is to evaluate that ICT-supported community-based charity platform. This usually involves user testing, gathering feedback and evaluating key performance indicators. According to the collected data some of them gave feedback and suggestions to improve this. And also, when testing the developed system used different browsers as well.

A. *Future recommendations*

There are new features can be added accordingly in future. Currently there is no place to add beneficiary's details. So, should add a place for that. Then, people can directly contact them.

And should establish and maintain mechanisms for ongoing user engagement and feedback collection. Regular feedback should be collected from donors, volunteers and beneficiaries to ensure the platform is responsive to evolving needs and expectations. A new feature needs to be developed for feedback control and review.

Also, must be updated with emerging technologies such as artificial intelligence (AI) and blockchain their potential applicability to charity-based online communities and investigate how these technologies can enhance user experiences and impact. Therefore, a 24/7 hours working chatbot should be developed to provide a better user experience.

Must be registered with NGOs. Establish partnerships with charitable organizations, NGOs and philanthropic institutions. Collaborate closely with these stakeholders to implement and refine the platform in real-world contexts. Use their expertise and networks to maximize impact on the platform.

Consider expanding the platform's reach beyond regional boundaries to address global charitable needs. Should adapt the platform to accommodate different languages like Sinhala and Tamil and facilitating international participation. Should be able to register with passport instead of National ID.

VI. Conclusion

Research on the community-based charity platform highlights its potential to bring about positive change. By connecting donors, volunteers and beneficiaries, it creates a platform for meaningful contributions. This research aims to improve charity and effectiveness with an emphasis on community engagement. Traditional charity models face challenges, but community-based solutions offer new possibilities. Transparency, user engagement, many projects and community involvement are key. This research provides insights for a more effective future in philanthropic communities.

Acknowledgment

I want to express my heartfelt gratitude to the people and organizations who made contributions to the success of this research project. Their immense support and dedication enabled me to complete the research successfully. I am grateful to Dr. Mohamed Sapraz, my research supervisor. He provided full support and assistance to my research every day. His guidance, expertise and unwavering support helped to complete this research. Another is my adored family. They provided the most physical and mental support. Also, I must thank all my colleagues who helped me to carry out the research successfully. I express my appreciation to the all who generously shared their time and insights, thereby contributing to the project's success.

References

[1] Gottlieb, B. H., & Bergen, A. E. (2010). Social support concepts and measures. In Journal of Psychosomatic Research (Vol. 69, Issue 5, pp. 511–520). Https://doi.org/10.1016/j.jpsychores.2009.10.001

[2] Nelson, D. L., & Campbell Quick, J. (1991). Social support and newcomer adjustment in organizations: Attachment theory at work? In journal of organizational behavior (vol. 12).

[3] Careemdeen, j. D. (2022). Level of community environmental support for student learning among secondary school children in sri lanka. Muallim Journal of Social Science and Humanities, 31–37. https://doi.org/10.33306/mjssh/190

[4] Chen, H., Li, W., Lyu, T., & Zheng, X. (2021). Understanding people's participation in online charities: a dual-process approach of trust and empathic concern. Industrial Management and Data Systems, 121(7), 1642–1663. https://doi.org/10.1108/IMDS-09-2020-0513

[5] Laily, I. L., Komarudin, O., Fadhilah, S., & Azurat, A. (n.d.). Progressive Learning Design Strategy to Improve Impact Maturity of Charity organizations.

[6] Wijesundara, R., & Aadhil, M. (n.d.). Development of a Web-Based Charity Organizations and Donation Management System: A Case Study. https://www.researchgate.net/publication/371575052

[7] Hackler, D., & Saxton, G. D. (2006). Sharpe) and has published work in Urban Affairs Review, Journal of Urban Affairs. In American Behavioral Scientist, and Annals of Cases on Information Technology. http://www.GregorySaxton.net/papers.html,

[8] Aidarkhan, A., Salamat, A., & Issin, N. (n.d.). Development of a crowdfunding web-platform to improve the effectiveness of medical fundraising for a charity fund.

[9] Saleh, H., Avdoshin, S., & Dzhonov, A. (2019). Platform for tracking donations of charitable foundations based on blockchain technology. Proceedings - 2019 Actual Problems of Systems and Software Engineering, APSSE 2019, 182–187. https://doi.org/10.1109/APSSE47353.2019.00031

Rubber Care - A Mobile Application to Detect the Diseases of Rubber Cultivation in Sri Lanka

Sewmini Asmadale
Faculty of Computing
NSBM Green University, Sri Lanka
sewminiasmadale@gmail.com

Mohamed Sapraz
Faculty of Computing
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 analyse 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.

Iinterviews 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

Fig 2. Phytophthora

Fig 3. Pestalotiopsis Fig 4. Pestalotiopsis

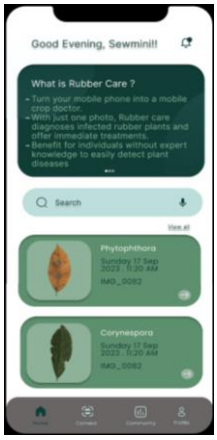


Fig 5. Landing Page Fig 6. Home Page

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.

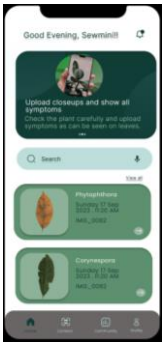


Fig 7. Home Page Fig 8. Home

Once an image is captured and incorporated into the scanning process, users are presented with a dedicated disease decryption page within the Resea n. This speci aluable 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.

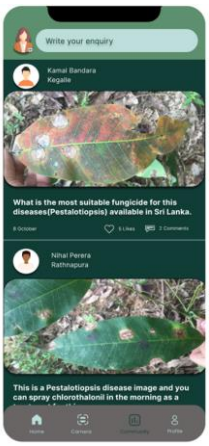


Fig 9. Community Page

Fig10.Recent Page

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.

ACKNOWLEDGMENTS

I wish to extend my heartfelt appreciation to all who played a vital role in the completion of this research project. My sincere thanks go to my dedicated supervisor, lecturer Mr.Dilhara Batanarachchige, and our final year research module coordinator, Head of the Department of Information and System Science senior lecturer Dr. Mohamed Sapraz for their unwavering guidance. I'm grateful to my family for their constant

encouragement, to my friends for their unwavering support. Special thanks to the participants who generously shared their insights. Each of you has made a significant contribution to the success of this research project.

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].