EduGuard – Mobile Application for Sex Education and Sexual Harassment Reporting for Sri Lanka

Sithmi Udapolawatta
Faculty of Computing
Sri Lanka Institute of Information Technology
Malabe, Sri Lanka
it20253158@my.sliit.lk

Nimesha Perera
Faculty of Computing
Sri Lanka Institute of Information Technology
Malabe, Sri Lanka
it20013264@my.sliit.lk

Kalya Perera
Faculty of Computing
Sri Lanka Institute of Information Technology
Malabe, Sri Lanka
it20110734@my.sliit.lk

Sajith Henarath
Faculty of Computing
Sri Lanka Institute of Information Technology
Malabe, Sri Lanka
it20025212@my.sliit.lk

Uthpala Samarakoon
Faculty of Computing
Sri Lanka Institute of Information Technology
Malabe, Sri Lanka
uthpala.s@sliit.lk

Dinithi Pandithage
Faculty of Computing
Sri Lanka Institute of Information Technology
Malabe, Sri Lanka
dinithipandithage@gmail.com

2024 6th International Conference on Advancements in Computing (ICAC) | 979-8-3315-1787-8/24/$31.00 ©2024 IEEE | DOI: 10.1109/ICAC64487.2024.10850953

***Abstract*—The lack of comprehensive sex education for Sri Lankans and inadequate systems for reporting sexual harassment are significant problems in the current society. These issues lead to limited awareness, compromised safety, and poor health outcomes. Many individuals lack access to accurate information and resources, exacerbating the risks associated with sexual health and safety. Addressing these gaps is crucial for fostering a well-informed and secure community. In this research, we introduce Eduguard, a mobile application aimed at improving sex education and sexual harassment reporting in Sri Lanka. This comprehensive approach addresses significant gaps in education and safety, promoting awareness and health through advanced technology. Eduguard features four key components: an age-appropriate sex education platform using face recognition, a sexual harassment reporting system utilizing voice recognition for discreet reporting, a blog and questionnaire platform with Natural Language Processing (NLP) for content filtering, and a sexually transmitted disease detection module employing image processing for early diagnosis. In conclusion, Eduguard represents a pioneering solution that combines education, safety, and health through the use of advanced technology. Its impact is substantial, offering a replicable model for regions facing similar challenges.**

***Keywords*— Image Processing, NLP, Sex Education, Sexual Harassment, Voice Recognition**

I. INTRODUCTION

Sex education is a critical aspect of overall educational development and mental health, yet it remains significantly underdeveloped in many third world and developing countries, including Sri Lanka. Unlike developed nations, where comprehensive frameworks for sexual education are integrated into the school curriculum, Sri Lanka faces a substantial gap in this area. For example, Scotland has successfully mainstreamed comprehensive sexuality education as part of a broader focus on health and well-being, placing it at the core of children's learning alongside literacy and numeracy [1]. In contrast, Sri Lanka's lack of structured sex education has contributed to severe conflicts between parents and their children regarding sexual relationships, which have been identified as potential triggers for suicidal behavior, especially among young people perceived to bring shame upon themselves and their families [2]. This crisis underscores the urgent need for effective sex education in Sri Lanka. To address this, the development of an age-based sex education platform using face recognition technology is proposed. This system aims to ensure that individuals access age-appropriate content, thereby fostering a more informed and healthier society.

Sexual harassment in Sri Lanka is a prevalent issue that impacts women in various settings, including the workplace, public transport, and online platforms. Studies have shown that workplace sexual harassment leads to adverse effects on personal lives, causing psychological, physical, and interpersonal problems [3]. Additionally, mobility issues for female travelers in Sri Lanka highlight the significant problem of sexual harassment while traveling, with a notable percentage of women facing harassment on buses and in public places [4]. Traditional methods for addressing this issue rely on self-reporting mechanisms and manual monitoring, which can be inadequate due to underreporting and delayed responses. Recent advancements in AI (Artificial Intelligence) and voice recognition technology present an innovative solution with the potential to enhance the detection and prevention of sexual harassment. This research paper explores the application of voice recognition technology as a proactive measure in preventing sexual harassment by leveraging Machine Learning (MI) capabilities to analyze and interpret vocal cues in real time.

In Sri Lanka, the internet has become a popular avenue for accessing information on sex education, with blogs and questionnaire platforms emerging as supposed educational resources. However, amidst their claims to offer valuable insights, many of these platforms contain inappropriate content, including explicit language and graphic imagery. This content is not suitable for users, particularly young people, as it can lead to misconceptions and negative attitudes towards sexuality. Additionally, the anonymity of online platforms can facilitate the spread of misinformation, further compromising the quality of sex education available online. To address this issue, we propose the use of Natural Language Processing (NLP) to filter content, ensuring that inappropriate words and imagery are detected and blocked when users post blogs and questionnaires. This approach aims to improve the quality and safety of online sex education resources, making them more suitable and informative for all users [5].

When considering sexual health aspects in Sri Lanka, addressing sexually transmitted diseases (STDs) is crucial, as they can spread through sexual contact and cause serious health issues if untreated. Common STDs, such as Chlamydia, Gonorrhea, Syphilis, HIV, and genital herpes, can be transmitted from mother to baby during pregnancy, childbirth, or breastfeeding. STDs like HIV and Hepatitis B & C (HBV & HCV) can also spread through sharing contaminated needles or syringes as well [6]. Misconceptions and cultural attitudes about STDs can make people feel isolated and hesitant to seek treatment, even with evident symptoms. Traditional diagnostic methods are often invasive and slow, causing further delays in treatment. However, recent advancements in medical imaging and computer technology such as Image processing can automatically analyze STD-related symptoms, aiding in quicker and more precise identification and diagnosis. The proposed application features a medical platform with STD advice, image processing to detect STD, and easy access to healthcare professionals and consultations [7].

II. LITRATURE REVIEW

The primary objective of this study is to develop EduGuard, an Android mobile application designed to address key challenges identified throughout the study. Integrating these distinct features into a single app represents an innovative approach not previously explored in Sri Lanka or globally.

A. Sex Education platform

Mobile learning transforms sex education by enabling accessible, flexible, and tailored experiences, essential for sensitive subjects like sex education. For instance, the "SexEdlk" YouTube channel provides free educational resources. However, its content lacks age-specific segmentation, making it less suitable for young viewers [8]. Studies highlight that mobile devices boost engagement and retention through interactive features. Tailoring content based on factors such as age and marital status effectively addresses users' developmental and relational needs. Privacy and security remain crucial for sensitive information in mobile applications, with methods like combining National Identity Card (NIC) and biometric verifications (selfie photo) strengthening user authentication [9]. Research reveals significant gaps in knowledge, attitudes, and skills in sexuality education across age groups in Sri Lanka. For example, knowledge levels span from 22% among ages 5-8 to 48% in the 12-15 age group, underscoring an urgent need for improvements in comprehensive sexuality education [10].

B. Sexual Harassment Prevention System

Various studies have explored the use of technology to prevent sexual harassment. Approaches include, Karomah et al. developing systems integrated with sexual education [11], Nagar et al. utilizing lexicon based and ML models to detect harassment in social media [12] , Angelin er al. implementing GPS based monitoring devices to send real time alerts to contacts or authorities in dangerous situations. [13]. These diverse technological solutions aim to enhance safety, raise awareness, and empower individuals to combat sexual harassment in various settings. This study aims to delve into the development of a mobile application prevent sexual harassment. Various research papers have proposed innovative methods to prevent sexual harassment using mobile applications. Putu et al. developed an interactive multimedia android-based application aimed at children aged 6-12 to educate them about sexual harassment and how to prevent it [14]. Bagas et al. developed an android application that allows victims to report incidents of sexual violence and harassment using their devices, providing reporting, counseling, and identification features [15]. Furthermore, the development of the HearMe mobile application includes unique features like lock screen access and an instant siren for quick assistance, enhancing the victim's ability to seek help effectively [16]. Rakshak is a project carried out by Delhi-based Bhart Vidyapeeth College of Engineering. They developed an android application that can detect speech patterns and generate SOS in case of an emergency [17]. The application uses audio snippets with speech commands requesting help or saying "stop" in distressed tones and detect distress in users, contacting emergency contacts. Circle of Six is a project developed for college students and their friends to stay close and prevent sexual harassments from happening. This application has a very user-friendly design and provides an easy to use and secure way to look out for each other [18]. These diverse approaches showcase the potential of mobile applications in combating sexual harassment and violence.

C. Blogs and Quesstionnaire Platform

Sri Lanka's cultural norms and societal pressures severely limit open dialogue on sensitive issues such as sex education, mental health, and gender equality. These constraints restrict access to accurate information, perpetuating social inequalities. Platforms like Xeducation aim to provide space for discussions, but they often struggle with inappropriate content, highlighting the need for effective content moderation to protect children and ensure the reliability of information. [19]. Balancing cultural sensitivities with open dialogue is essential for fostering social progress. Well-moderated platforms for sex education can encourage healthy discussions, provide accurate information, and create safe spaces for individuals to share their experiences, reducing inequalities. However, existing studies such as those by Gerald R. Gendron [20], Hitesh Kumar Sharma et al [21] focus mainly on technical aspects like machine learning models and NLP without addressing the cultural challenges in regions like Sri Lanka. This research bridges that gap by developing culturally adaptive NLP algorithms that enhance content moderation and provide context-sensitive questionnaire platforms, ensuring both accuracy and cultural appropriateness. [22].

D. STD Detection using Image Processing and Clinical Recommendation Algorithm

Various approaches have been taken globally to detect STDs. For example, Feller et al. utilized clinical notes and natural language processing (NLP) to assess HIV risk, developing predictive algorithms with acceptable performance [23]. Xu et al. developed machine learning models and a web-based risk-prediction tool using consultation data from individuals re-tested for STDs, predicting their future risk over a 12-month period. [24]. González-Alday et al. developed a deep learning system for classifying genital lesions caused by herpes, warts, and condylomas using convolutional neural network integrated into a web framework. [25]. Platforms like "STDcheck" provide consultation facilities, test kits, and clinic locators, but are limited to the United States [26]. "know4sure" offers HIV risk analysis and clinic recommendations specific to Sri Lanka, using machine learning models [27].

While these studies demonstrate the use of predictive models or classification systems, they are often limited to specific regions or specific STDs, and none of them apply image processing for broader STD detection in the Sri Lankan context. Existing approaches mostly focus on HIV risk prediction or are tailored to different healthcare systems. Our research addresses this gap by applying image processing techniques to detect STDs such as Syphilis, Gonorrhea and Healthy conditions, providing a novel solution specifically designed for the Sri Lankan community.

III. METHODOLOGY

The bellow figure 1 represents the overall system architecture diagram of the proposed application.

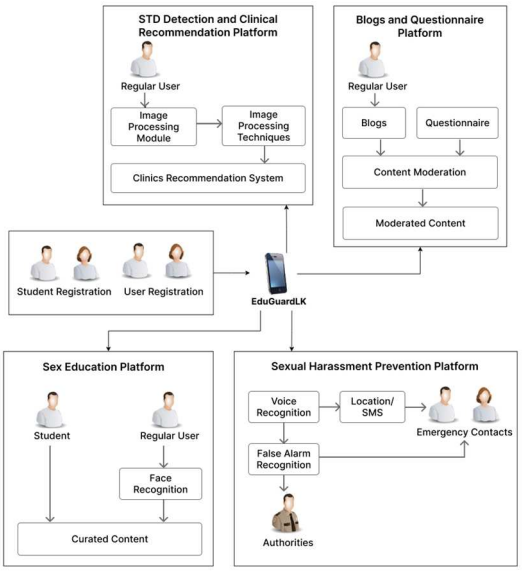

Fig. 1. Overall System Architecture

A. Data Collection Survey

A data collection survey was designed to gather insights into public perceptions and preferences regarding sex education platforms. The survey was administered to a diverse sample of participants across different age groups, genders, and socio-economic backgrounds. Participants were recruited from various sources, including workplaces, educational institutions, schools, and social media connections. Informed consent was obtained from all participants prior to their participation in the survey. Data collection was conducted anonymously to ensure participant confidentiality and minimize response bias.

According to the survey results, a significant portion of participants have not received formal sex education in school. Also, the majority have encountered or observed physical harassment and have bad experiences in blogs and questionnaire platforms as well. Furthermore, the data reveals a gap in knowledge regarding STDs among people in Sri Lanka.

B. Sex education platform

This component aims to provide tailored sex education content for everyone. The content is customized based on the user's age and marital status to ensure relevance. For example, a 33-year-old married man would receive content focused on managing a healthy marital life, which would not be suitable for a teenager. To achieve this customization, users are required to register with their National Identity Card (NIC) and a selfie taken with their phone camera. A machine learning algorithm will detect the user's age and verify that the selfie matches the NIC photo to confirm they are the same person. If the user is over 18 years old and their identity is verified, the app will provide learning content appropriate for their age. For users under 18, who are considered students, there is a special registration process. They can get an access token from their school, whether government or private, which will be used during registration. This token ensures they receive age-appropriate learning content. The system is designed to collaborate with schools to distribute these access tokens effectively.

The backend API is built using Python and Flask, ensuring a robust and scalable server-side framework. The system uses OpenCV and the face-recognition library to process and verify the user's selfie against their National Identity Card (NIC) photo, employing advanced image recognition techniques. The face-recognition models and dlib are integral in accurately detecting and comparing facial features. Additionally, Tesseract OCR is used to extract text information from the NIC for further verification. By combining these technologies, the application ensures that the user's identity is authenticated accurately and efficiently, allowing for the delivery of age-appropriate sex education content tailored to each user's needs. This approach not only enhances security but also personalizes the learning experience, making it more relevant and effective.

C. Sexual harassment prevention system

The system architecture of this component is meticulously designed to ensure an effective and responsive emergency SOS system. The user interface is designed to be intuitive, allowing users to easily activate the SOS system through either a wake word or touch input. The wake word feature is enabled by integrating the Picovoice module, which continuously listens for a predefined trigger word. The porcupine wake word detection engine by picovoice is used to detect the wake word, which is configured to operate efficiently in the background, continuously listening for the wake word with sensitivity settings tuned to balance responsiveness and minimize false activations. Upon detection, the application seamlessly transitions into SOS state. Alternatively, users can manually activate the SOS state via a dedicated button within the app, providing an essential alternative for situations where the wake word might not be feasible.

A Speech Emotion Recognition (SER) model is embedded within the application to monitor the users emotional state during the SOS state. This model analyzes audio inputs in real time, detecting distress related emotions such as fear, anger, and sadness. The decision logic processes the SER model's output and validates the SOS state activation, reducing the likelihood of false alarms. If a detected emotion does not indicate distress, the system prompts the user for confirmation before proceeding with emergency actions. The development of the SER model involved the use of multiple datasets to ensure a robust and comprehensive training process. Ryerson Audio Visual Database of Emotional Speech and Song (RAVDESS), Crowd –Sourced Emotional Multimodal Actors Dataset (CREMA-D), Toronto Emotional Speech Set (TESS), and Surrey Audio Visual Expressed Emotion (SAVEE) datasets were used to ensure the diversity and the comprehensive representation of the data. Each of these

datasets offers a unique contribution to the understanding of emotional expression through speech.

Collectively, these datasets represent a diverse array of emotional expressions, captured through various contexts, demographics, and performance styles. The combined dataset exhibits a balanced distribution of emotions, with each emotion category represented by a similar number of instances. The emotions included in the dataset are happy, sad, angry, fear, disgust, neutral, and surprise. Each emotion category comprises a substantial number of instances, with approximately 1923 instances. However, the neutral and surprise emotion categories have slightly fewer instances, respectively 1895 and 652 occurrences, compared to other emotions. Despite these minor imbalances, the overall distribution of emotions in the dataset remains uniform.

The process of feature extraction was done using fundamental audio features such as the Zero Crossing Rate (ZCR), Root Mean Square Energy (RMSE), and Mel Frequency Cepstral Coefficients (MFCCs). Also, audio shifting techniques are used to introduce variations in the audio data, thereby enhancing the feature set's robustness. These techniques include the addition of noise and the alternation of pitch. The introduction of noise introduces randomness and variability, simulating real world conditions where audio recordings may contain background noise or interference. Similarly, altering the pitch enables the model to learn from a wider range of pitch variations, improving its ability to generalize across different vocal expressions.

Following the feature extraction, the combined dataset was split into training and test sets. The training set was allocated 80% of the combined dataset while the other 20% was allocated for the test set. The SER model was built upon a sequential one-dimensional Convolutional Neural Network (CNN1D), which achieved high accuracy, through rigorous training and a validation process. The model's performance was evaluated periodically on a validation set to guide the training process and prevent overfitting. Metrics such as accuracy, precision, recall and f1 score were used to assess the model performance and its effectiveness in recognizing emotional expressions from audio data.

D. Detecting Inappropriate word content using Machine Learning

In Sex Education is Crucial in providing young people with accurate information about sexuality relationship and reproductive health however ensuring that online platforms for sex education, such as blog and questionnaire, remain free from inappropriate content is a significant challenge. Inappropriate content can include misinformation, explicit material, or harmful advice that may negatively impact the audience, particularly young individuals seeking reliable information.

The methodology begins with finding dataset. It found from Kaggle. Then importing various libraries essential for data processing, analysis, and machine learning tasks, including pandas and numpy for data manipulation, CountVectorizer for text-to-numerical conversion, and train_test_split for dataset splitting. Additional libraries such as matplotlib.pyplot, seaborn, WordCloud, re, nltk, warnings, string, and stopwords are used for data visualization and text processing. The dataset, stored in a CSV file named "labeled_data.csv," is loaded into a DataFrame called df, and its structure is previewed using print(df.head()). A new column named labels is created to map numerical values in the class column to descriptive strings, enhancing data interpretability. A clean() function is defined to preprocess the text data by converting it to lowercase and removing URLs, HTML tags, punctuation, digits, and stopwords. This function is applied to the "tweet" column, resulting in cleaned text data. The cleaned text is then transformed into a numerical format using CountVectorizer, creating a matrix of token counts where each row represents a document and each column a unique word. The dataset is split into training and testing sets using train_test_split, with 33% reserved for testing and 67% for training. Finally, a DecisionTreeClassifier is initialized and trained on the training data, making it ready to predict new, unseen data.

E. Detecting STD using Image Processing

The process of obtaining a suitable dataset for STD detection component involved extensive exploration of various medical databases and websites, including STIAtlas, DermNet, and Devon Sexual Health [28, 29, 30]. Given the sensitivity of the subject matter, the unavailability of pre-defined datasets necessitated the creation of a custom dataset. This dataset was meticulously curated to include a diverse range of images depicting both healthy and infected body parts relevant to STD infections. Careful categorization ensured that the dataset encompassed a comprehensive representation of visual data related to different types and severities of STDs. Preprocessing techniques such as image resizing, rescaling, and data augmentation were employed to enhance dataset quality and diversity, thus optimizing the effectiveness of subsequent model training and evaluation.

Following the acquisition and preprocessing of the custom dataset, the next step involved the division of the dataset into distinct subsets for training, testing, and validation purposes. The training dataset, comprising 80% of the total dataset, served as the primary source for training the STD detection model. Subsequently, a Convolutional Neural Network (CNN) model was selected for its proven effectiveness in image classification tasks, particularly in the context of medical image analysis. The CNN model underwent rigorous training using the training dataset, with parameters optimized to maximize model performance. To ensure the generalizability of the trained model, evaluation was conducted using both the testing and validation datasets. Performance metrics such as accuracy and loss were employed to quantitatively assess the model's ability to accurately distinguish between healthy and infected body images, providing insights into its overall effectiveness and reliability.

The implementation of the trained CNN model was realized through the utilization of industry-standard tools and libraries, including Jupyter Notebook, Python programming language, TensorFlow framework, and Keras deep learning library. These resources provided a robust and flexible development environment for building, testing, and refining the STD detection model.

IV. RESULTS & DISCUSSION

The following figure 2 represents main interfaces of EduGuard Application.

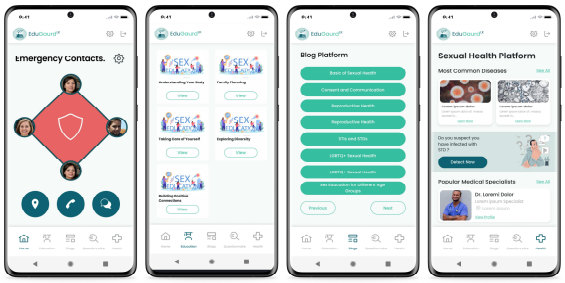

Fig. 2. Interfaces of EduGuard application

The sex education platform successfully provided tailored sex education to various user groups in Sri Lanka. Using machine learning and image processing, it accurately verified user identities and ages, delivering appropriate content. Adults received customized information, while students accessed content through secure tokens from their schools. Initial trials showed high user satisfaction, improved understanding of sexual health, and a strong sense of security. Collaboration with schools ensured broad access, effectively bridging the gap in sex education and aligning with global best practices.

Compared to baseline models such as Support Vector Machine (SVM), LSTM and previous approaches in SER, the CNN1D model shows significant improvements in accuracy, expressing the CNN1D model's ability to capture temporal and spectral features has resulted in superior performance. As shown in figure 3, the overall accuracy of the model is 93%, indicating its high effectiveness in correctly identifying emotional states from audio data. The model performed exceptionally well in detecting neutral and surprise emotions, as indicated by the highest F1 score of 0.94. However, there is a slight dip in recall for the angry emotion, suggesting some misclassifications that need further investigation. The balanced performance across most emotions highlights the model's effectiveness, but further tuning may be needed to improve recall for certain categories.

	precision	recall	f1-score	support
angry	0.98	0.89	0.93	1531
disgust	0.91	0.94	0.92	1549
fear	0.91	0.92	0.91	1552
happy	0.90	0.93	0.92	1516
neutral	0.96	0.92	0.94	1557
sad	0.91	0.94	0.93	1520
surprise	0.91	0.96	0.94	505
accuracy			0.93	9730
macro avg	0.93	0.93	0.93	9730
weighted avg	0.93	0.93	0.93	9730

Fig. 3. SER Model Accuracy

Future research could focus on expanding the dataset to include more diverse emotional expressions and demographic variations. Additionally, exploring multimodal approaches that combine audio with environmental contexts could further enhance emotion recognition accuracy. Also, continuous learning and adaptation mechanisms can also be integrated to improve the model's performance over time as more data becomes available.

Detecting hate speech in sex education content, such as blogs and questionnaires, involves several steps including data collection, preprocessing, feature extraction, model training, and evaluation. Initially, relevant text data is gathered from blogs and survey responses. This data undergoes preprocessing to remove unnecessary elements like HTML tags, URLs, and punctuation, followed by tokenization and lowercasing to ensure consistency. Feature extraction techniques such as TF-IDF (Term Frequency-Inverse Document Frequency) are then used to represent the text data in a meaningful way. Various machine learning models like Logistic Regression or deep learning models like BERT can be trained on this labeled data to identify hate speech. For instance, in a simplified implementation using Logistic Regression and TF-IDF on a small dataset, the model achieved high performance, demonstrating effectiveness in detecting hate speech with metrics such as perfect accuracy, precision, recall, and F1 scores as in figure 4. However, for practical applications, a larger and more diverse dataset is necessary to train a robust model. Ethical considerations are also crucial when handling sensitive topics like sex education and hate speech, ensuring data privacy and understanding the impact of false positives and negatives.

Fig. 4: Accuracy of Hate speech recognition model

According to figure 5, the image processing module demonstrated strong performance in detecting sexually transmitted diseases, achieving an overall accuracy of 87%. These results confirm that image processing is an effective method for detecting STDs, highlighting the potential of computer vision to provide a cost-effective, time-efficient alternative to traditional diagnostic methods. The novelty of this approach lies in enabling individuals to identify STD symptoms without relying on costly and time-consuming test kits, and in facilitating the recommendation of suitable clinicians as required, thus enhancing privacy. However, the variability in image quality and resolution, stemming from multiple sources, might have influenced the outcomes. Additionally, the study's limited dataset size suggests that incorporating more images could improve robustness and generalizability. Future research should focus on standardizing image quality and expanding the dataset to further validate and enhance the model's performance.

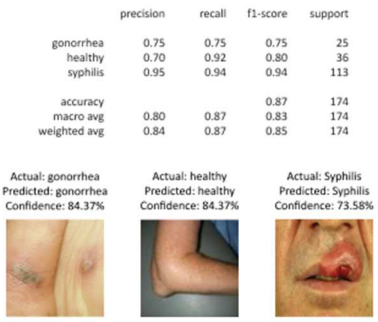

Fig. 5. STD Detection model Accuracy and predicted labels

In comparison with existing systems, EduGuard's sex education platform addresses limitations like those in

"SexEdlk" by offering age-appropriate, secure content verified through machine learning. The sexual harassment prevention system enhances features seen in tools like "HearMe" and "Rakshak," adding real-time reporting and voice recognition for distress detection. The blogs and questionnaire platform leverages NLP for safe, moderated discussions, surpassing platforms like "Quora" and "Xeducation" in addressing sensitive topics. Unlike platforms such as "STDcheck" and "know4sure", EduGuard uses image processing to provide a more accessible, cost-effective solution for STD detection. These results highlight EduGuard's significance that effectively bridges critical gaps in healthcare, safety, and education.

V. Conclusion

In this study, we proposed EduGuard, an innovative Android application integrating a sex education platform using face recognition, a harassment reporting system with voice recognition, a blog and questionnaire platform employing natural language processing, and a STD detection module using image processing. Findings indicate that the system provides a comprehensive solution by combining these technologies to offer an accessible and effective platform with productive results. EduGuard has the potential to enhance sexual health education and support systems in Sri Lanka. Furthermore, the application's innovative approach can serve as a model for similar initiatives in other regions.

However, the current version of the application is limited to the English language and the Android platform, which restricts its accessibility to a wider audience. Future research should focus on enhancing the accessibility and functionality of the system. This includes implementing the application in the Sinhala language to reach a broader audience and developing an iOS version for compatibility across various mobile platforms. By continuously developing and refining EduGuard, we aim to empower individuals to take control of their sexual education and safety, thereby reducing the stigma surrounding discussions on sexual health and contributing to the overall well-being of the community.

References

[1] UNESCO, "Comprehensive Sexuality Education Implementation Toolkit," [Online]. Available: https://csetoolkit.unesco.org/toolkit/getting-started.

[2] G. Crowley, P. Bandara, L. Senarathna and A. Malalagama, "Sex education and self-poisoning in Sri Lanka: an explorative analysis," *BMC Public Health,* January 2022.

[3] A. N. Mingappulige and C. De Alwis, "Breaking the Silence: Exploring the Prevalence and Effects of Sexual Harassments in Sri Lanka's Workplaces.," *Vallis Aurea,* vol. 9, no. 1, pp. 5-18, 2023.

[4] T. Karunarathne and N. Amarasingha, "Travel Issues of Sri Lankan Females," *Proceedings of the SLIIT International Conference On Engineering and Technology,* vol. 1, 2022.

[5] A. Balayn, J. Yang, Z. Szlavik and A. Bozzon, "Automatic Identification of Harmful, Aggressive, Abusive, and Offensive Language on the Web: A Survey of Technical Biases Informed by Psychology Literature," *ACM Transactions on Social Computing,* vol. 4, no. 3, 2021.

[6] P. C. Adamson, M. J. Loeffelholz and J. D. Klausner, "Point-of-Care Testing for Sexually Transmitted Infections: A Review of Recent Developments," *College of American Pathologists,* 18 August 2020.

[7] T. Spence, I. Kander, J. Walsh, F. Griffiths and J. Ross, "Perceptions and Experiences of Internet-Based Testing for Sexually Transmitted Infections: Systematic Review and Synthesis of Qualitative Research," *Journel of Medical Internet Research,* 2020.

[8] P. Premadasa, "SexEdLk," [Online]. Available: https://www.youtube.com/@SexEdLk.

[9] A. K. Jain, P. Flynn and A. A. Ross, Handbook of biometrics, Springer Science & Business Media, 2007.

[10] M. Jayantha and C. Chandrakumara, "Investigating the Comprehensiveness of Sexuality Education Curriculum in Sri Lanka," in *2nd International Research Symposium - 2022*, 2022.

[11] A. Karomah, "Preventing Sexual Violence through AI-Based Smart Hug Website in Educational Institutions in Semarang," *BICC Proceedings,* vol. 1, pp. 59-66, 2023.

[12] A. Nagar, M. Bhat, K. Sneha Priya and K. Rajeshwari, "A Holistic Study on Approaches to Prevent Sexual Harassment on Twitter.," in *Machine Learning for Predictive Analysis. Lecture Notes in Networks and Systems, vol 141. Springer,*, Singapore, 2021.

[13] T. Angelin, Monisha, Sharean and M. Natalie, "Smart wearable gadget for women safety.," *IIP Series,* vol. 3, pp. 188-203, 2024.

[14] D. Putu , Yudhi, Ardiana and L. Hendrika, Loekito, "Perancangan multimedia interaktif untuk materi perlindungan anak terhadap pelecehan seksual berbasis mobile.," *SENADA (Seminar Nasional Desain dan Arsitektur),* vol. 2, pp. 216-222, 2019.

[15] A. Bagas and E. Supraptono, "The Development of an Android-Based "LaporKPS" Application to Support the Service Center for Reports of Sexual Violence and Harassment Cases.," *Data Science: Journal of Computing and Applied Informatics,* vol. 6, no. 2, pp. 111-124, 2022.

[16] A. S. Ahmed, M. Al-Zihad, T. Adhikary, M. A. Razzaque and A. Sharmin, "HearMe: A smart mobile application for mitigating women harassment," in *2016 IEEE International WIE Conference on Electrical and Computer Engineering (WIECON-ECE)*, 2016-12.

[17] D. C. Sharma, "DowntoEarth: An app that can detect distress speech and trigger SOS," 21 October 2019. [Online]. Available: https://www.downtoearth.org.in/news/science-technology/an-app-that-can-detect-distress-speech-and-trigger-sos-67354.

[18] D. Levine, N. Schwartzman, C. C. Moran and T. Cabus, "Devpost: Circle of 6," 15 October 2011. [Online]. Available: https://appsagainstabuse.devpost.com/submissions/4900-circle-of-6.

[19] "XeducationSL," 2022. [Online]. Available: https://www.xeducationsl.com/.

[20] G. R. Gendron, "Natural Language Processing: A Model to Predict a Sequence of Words," in *MODSIM World 2015*, Norfolk, 2015.

[21] H. K. Sharma, K. Kshitiz and Shailendra., "NLP and Machine Learning Techniques for Detecting Insulting Comments on Social Networking Platforms," in *2018 International Conference on Advances in Computing and Communication Engineering (ICACCE)*, Paris, 2018.

[22] "Quora," [Online]. Available: https://www.quora.com/Is-there-an-18-section-of-Quora.

[23] D. J. Feller, J. Zucker, M. T. Yin, P. Gordon and N. Elhadad, "Using Clinical Notes and Natural Language Processing for Automated HIV Risk Assessment," *Journal of Acquired Immune Deficiency Syndromes,* 2018.

[24] X. Xu, Z. Ge, E. Chow, Z. Yu, D. Lee, J. Wu, J. Ong, C. Fairley and Zhang, "A Machine-Learning-Based Risk-Prediction Tool for HIV and Sexually Transmitted Infections Acquisition over the Next 12 Months," *Journel of Clinical Medicine,* 2022.

[25] R. González-Alday, F. Peinado, D. Carrillo and V. Maojo, "Deep Learning for Clinical Image Classification of Genital Lesions Caused by Sexually Transmitted Diseases," in *Royal National Academy of Medicine*, 2023.

[26] "STDcheck," [Online]. Available: https://www.stdcheck.com/.

[27] "Know4sure," [Online]. Available: https://know4sure.lk/.

[28] Melbourne Sexual Health Centre, "STI Atlas," [Online]. Available: https://stiatlas.org/.

[29] DermNet, [Online]. Available: https://dermnetnz.org/.

[30] Devon Sexual Health, [Online]. Available: https://www.devonsexualhealth.nhs.uk/professionals/image-gallery/.