Computing, Information Systems, Development Informatics & Allied Research Journal
Vol. 14 No. 3, September 2023 - https://www.isteams.net/cisdijournal

A Synchronized Mobile Responsive Application for Suicide Ideation Detection

[1]Ezea, I. L, [2]Eneh, A. H., [3]Agu, N. M. & [4]Bakpo, F.S.
[1,2,3,4]Department of Computer Science
Faculty of Physical Sciences
University of Nigeria, Nsukka.
E-mail: [1]ezea.ikenna@funai.edu.ng, [2]agozie.eneh@unn.edu.ng, [3]monica.agu@unn.edu.ng, [4]francis.bakpo@unn.edu.ng
Phone: +2348025107142

ABSTRACT

The increasing concern over mental health issues, particularly suicide ideation, necessitates innovative and accessible solutions for early detection and intervention. This paper presents a synchronized mobile responsive application designed to detect suicide ideation efficiently and effectively. Leveraging the ubiquity of smartphones and the power of real-time data processing, our application offers a user-friendly interface that facilitates seamless engagement while continuously monitoring users' online behavior and communications. The system integrates advanced natural language processing, machine learning algorithms, and user behavioral analysis to identify potential signs of suicide ideation. By synchronizing data across multiple platforms (social network and academic applicataion) and providing timely alerts to both users and relevant support networks, the application aims to bridge the gap between technology and mental health care. This novel approach holds promise in enhancing suicide prevention efforts by providing a proactive and real-time tool for early detection and intervention, ultimately saving lives and promoting mental well-being.

Keywords: Synchronized, Mobile Responsive, Application, Suicide Ideation, Detection

CISDI Journal Reference Format
Ezea, I. L, Eneh, A. H., Agu, N. M. & Bakpo, F.S. (2023): A Synchronized Mobile Responsive Application for Suicide Ideation Detection. Computing, Information Systems, Development Informatics &Allied Research Journal. Vol 14 No 3, Pp 43-56. Available online at https://www.isteams.net/cisdijournal. dx.doi.org/10.22624/AIMS/CISDI/V14N3P4

1. INTRODUCTION

In an era dominated by digital connectivity, the intersection of technology and mental health has gained significant attention. One pressing issue within this domain is the detection of suicide ideation an alarming concern that requires innovative, accessible, and real-time solutions. Suicide remains a global public health challenge, and early identification of individuals at risk is critical for prevention and intervention. As a response to this challenge, we introduce a synchronized mobile responsive application designed to detect suicide ideation effectively. The alarming rise in suicide rates, particularly among undergraduates and vulnerable populations, has prompted a growing emphasis on leveraging technology to provide timely support. Suicide ideation, often the precursor to more severe actions, is increasingly expressed through online interactions and social media platforms. This shift presents both an opportunity and a challenge; an opportunity to use digital data for early detection, and a challenge in sifting through vast amounts of online content to identify at-risk individuals.

This article discusses a novel approach to suicide ideation detection namely, a mobile responsive application that synchronizes data from multiple platforms, offering continuous monitoring, real-time analysis, and immediate alerts when potential suicide ideation is detected. By taking advantage of the widespread use of smartphones, this application provides users with an unobtrusive and user-friendly means of seeking help and support when they need it most. Furthermore, it assists caregivers, mental health professionals, and support networks in identifying and reaching out to individuals who may be in distress. To address this critical issue, our application integrates state-of-the-art technologies, including natural language processing, machine learning, and behavioral analysis, to scrutinize users' online activities and communications. The aim is not only to detect potential suicide ideation but also to do so proactively and in real-time. The potential impact of such a tool is far-reaching, as it could save lives, alleviate the burden on mental health professionals, and contribute to a better understanding of the evolving landscape of mental health in the digital age.

In the sections that follow, we will delve into the components, functionalities, and the methodology behind our synchronized mobile responsive application. We will also discuss the ethical considerations, potential challenges, and future implications of this technology. Ultimately, our objective is to introduce a comprehensive solution that not only addresses the immediate issue of suicide ideation detection but also raises broader questions about the evolving role of technology in mental health support and intervention.

2. LITERATURE REVIEW

The detection of suicide ideation in the digital age has become an imperative endeavor, and researchers have employed various methodologies to address this critical issue. This literature review provides an overview of key studies that have contributed to the understanding and advancement of suicide ideation detection, culminating in the development of a synchronized mobile responsive application for this purpose. Ezea's work [1] is among the authors that have made efforts in providing mathematical model for undergraduate suicide ideation detection. His model provided a foundation for subsequent research in this field, highlighting the importance of mathematical and computational sciences in addressing this complex problem. This mathematical approach paved the way for a more data-driven and systematic exploration of suicide ideation detection.

In another work he [2] extended his research in conducting a systematic review of machine learning classification approaches for suicide ideation detection. This study emphasized the significance of machine learning techniques in automating the detection process, offering a structured and data-driven means of identifying at-risk individuals. The systematic review underlined the need for robust and effective classification models. The emergence of online social networks as platforms for expressing suicide ideation has been a focal point of several studies. Masuda et al. [3] delved into understanding the suicide ideation of individuals within online social networks, shedding light on the dynamics of expressing suicidal thoughts in digital spaces. Jashinsky et al. [4] extended this exploration by tracking suicide risk factors through Twitter, emphasizing the potential of social media data for detection. Twitter, as a prominent social media platform, has been a primary focus of research.

Varathan and Talib [5] proposed a suicide detection system based on Twitter, harnessing the wealth of data available on this platform. Gunn and Lester [7] analyzed Twitter postings in the 24 hours preceding a suicide, providing insights into linguistic patterns that may indicate impending risk. Analyzing microblogs, such as those found on Twitter, has been an effective avenue for understanding stress and suicidal content. Li et al. [8] focused on identifying stressful periods and stressor events among teenagers in microblogs, shedding light on the context and triggers of suicide ideation. Colombo et al. [9] explored the connectivity and communication patterns of suicidal users on Twitter, highlighting the propagation of such content.

The use of natural language processing (NLP) has become increasingly prominent in suicide ideation detection. Ji et al. [11] employed supervised learning for the detection of suicidal ideation in online user content, harnessing NLP techniques to discern explicit and implicit indicators. Cao et al. [12] introduced a latent suicide risk detection model on microblogs, utilizing suicide-oriented word embeddings and layered attention to enhance the accuracy of detection. Recognizing the temporal dynamics of suicidal content, Sawhney et al. [13] introduced a time-aware transformer-based model for suicide ideation detection on social media. This approach acknowledges the evolving nature of risk factors and sentiments expressed over time, enhancing the precision of detection.

The use of personal knowledge graphs has also gained attention. Cao et al. [14] proposed building and using personal knowledge graphs to improve suicidal ideation detection on social media. This innovative approach integrates individual-specific information to enhance the accuracy of detection. In summary, the literature reviewed here underscores the evolution of suicide ideation detection methods, from mathematical models and machine learning to the analysis of social media data and the incorporation of natural language processing and temporal considerations. Building upon this foundation, our synchronized mobile responsive application aims to provide a comprehensive, real-time solution to address the critical issue of suicide ideation detection in the digital age.

3. MATERIALS AND METHODS

In this section, we elucidate the materials, technologies, and methods employed in the development of our synchronized mobile responsive application for suicide ideation detection. The successful creation of such a critical tool hinges on the careful selection of materials and the application of robust methodologies. This section serves as a blueprint for the technical aspects of our research.

3.1. Data Collection and Preprocessing

- **Data Sources**: The primary data sources for our study encompassed popular social media platforms like Reddit Suicide and Depression Watch, X (formally Twitter), and public forums like Alex Ekwueme Federal University (AEFUNAI) Facebook Forum. These sources were chosen due to their widespread use for sharing personal experiences and emotions, including discussions related to suicide ideation.
- **Data Collection**: Data were collected using manual method, web scraping and application programming interfaces (APIs) provided by the respective platforms. This process spanned from February, 2022 to July, 2023, ensuring the inclusion of a wide range of temporal expressions of suicide ideation.
- **Data Preprocessing**: Raw data underwent a series of preprocessing steps to ensure uniformity and quality. This encompassed text cleaning, which involved removing special characters, formatting inconsistencies, and irrelevant symbols. Tokenization was applied to break down text into individual words and phrases, facilitating subsequent analysis. Additionally, sentiment analysis was conducted to assess the emotional tone of the text data, categorizing content as positive, negative, or neutral.
- **Anonymization**: To protect the privacy of individuals, all personally identifiable information, such as user names, specific locations, and contact details, was either removed or anonymized. This step was crucial to comply with ethical standards and legal regulations.

3.2. Responsive Application Development Method

The synchronized mobile application was developed for both Android and iOS platforms, ensuring broad accessibility. Technologies such as Python, MySQL were employed to facilitate cross-platform development.

3.2.1. System Specification

The project's execution involved the utilization of both hardware and software resources. A Hewlett Packard (HP) computer system served as the major implementation hardware, while a selection of Integrated Development Environments (IDEs) and application development libraries were employed for the coding process. Comprehensive information regarding the specific hardware and software specifications is provided in Table 1 and Table 2, respectively.

Table 5: Application Development Tools

Compoent	Specification
Operating System	Windows 10 Pro
Processor	Intel ® Pentium CPU N3710 @ 1.60GHz
Installed Memory (RAM)	4.00 GB
System Type	64-bit Operating System

Table 6: Computer System Specification

Component	Specification
Python IDE	Google Colab
Python Library	Pandas, Sklearn, Matplotlib, Numpy, Seaborn
Database System	Mysql 8.0
File System	MS Excel CSV
Web Development Tool	HTML, CSS, Javascript
HTTP Servers	Django

3.2.2 Synchronization Architecture

Architecture

To enable real-time data synchronization and monitoring, we developed a scalable and responsive architecture for our mobile application. This involved backend systems, databases, and APIs designed to process incoming data streams continuously. The architecture consists of students' application which is composed of social network and daily academic routing application. Social Network Conversations coming from the social network application goes through the database to Machine Learning Model for classification after which it is synchronized with students' academic activities before it being transmitted as an alert to the admin who may be mental health professional or any assigned personnel.

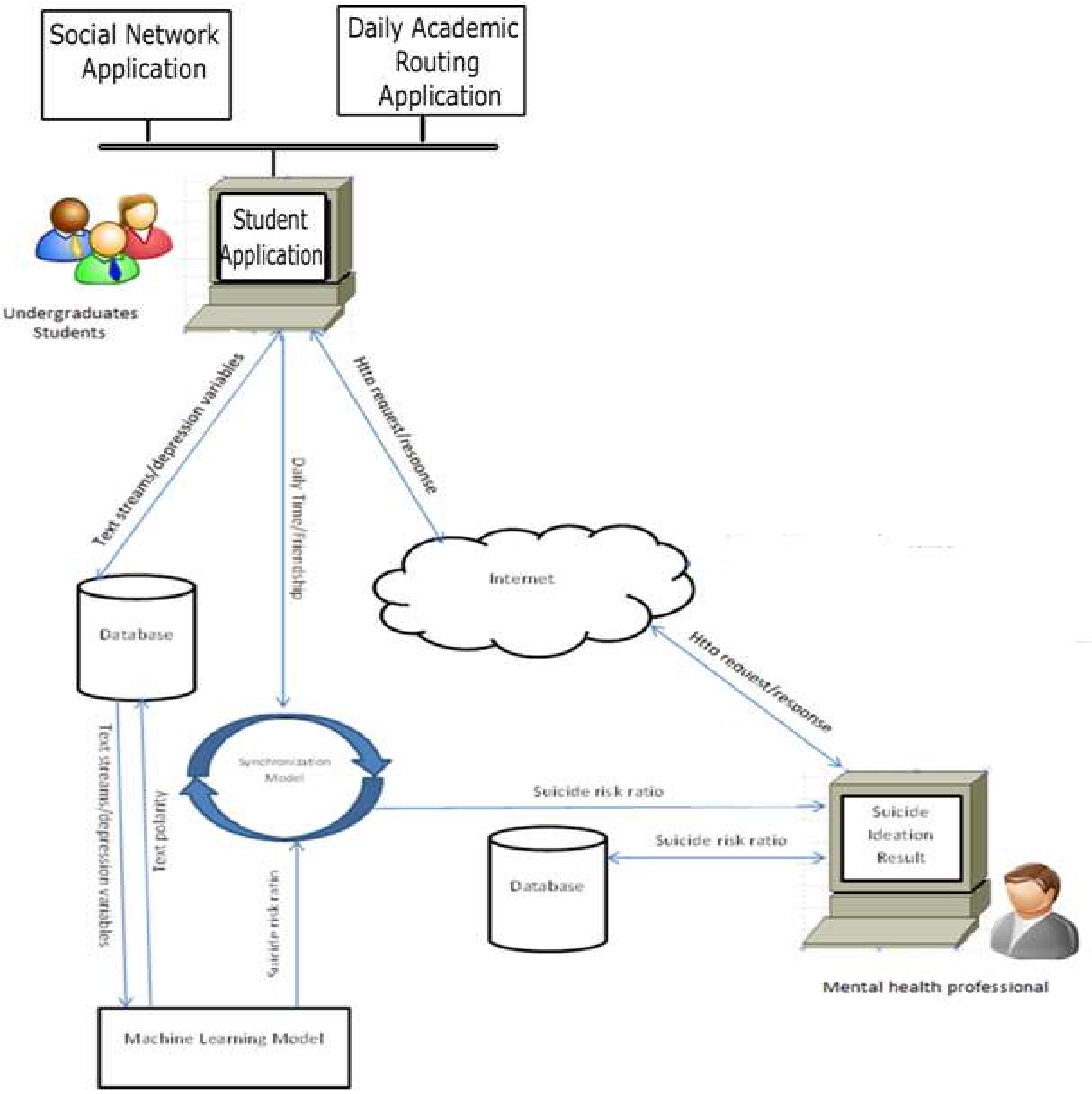

Figure 8: Application Architecture

3.3. Proposed System Framework

The proposed system framework (Figure 2) outlines a series of activities aimed at suicided ideation detection. The key phases involved in this framework are as follows:

- **Dataset Cleaning**: In this phase, the dataset is cleaned to eliminate inconsistencies, such as redundant and missing values. This ensures that the dataset is in a consistent state, which is crucial for model performance.
- **Feature Selection**: To improve classification performance, irrelevant and redundant features in the dataset are identified and removed.

- **Dataset Splitting**: The dataset is divided into training, evaluation and testing sets, typically in a 70:15:15 ratio. This division helps control bias and variance issues during model training and evaluation.
- **Training Classifier**: This phase involves adjusting the model's parameters to achieve a clear separation of target classes. This enhances the model's ability to make accurate predictions when given new data.
- **Testing**: The model is evaluated using test data to assess its performance. Evaluation metrics include accuracy and the creation of a confusion matrix.
- **Analysis**: The model's performance is further evaluated based on various metrics, including Receiver Operator Characteristics (ROC), Confusion Matrix, and Accuracy. The choice of the classifier is determined by how it performs against others using these metrics.
- **Prediction**: In this final stage, the system determines whether a given data sample indicates the presence of suicide ideation and provides a percentage likelihood of developing suicide.

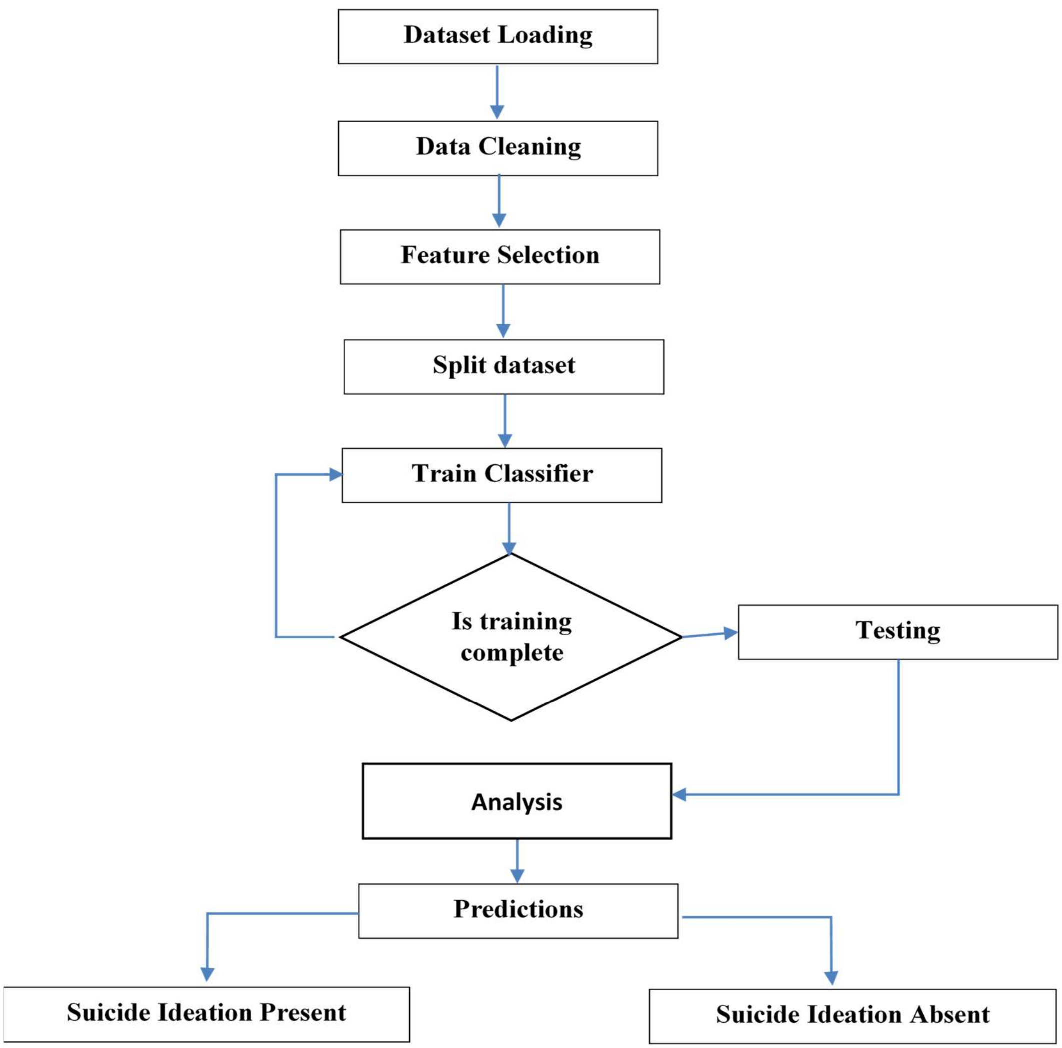

Figure 9: Suicide Ideation Detection Framework

Overall, this framework outlines a systematic approach to predict suicide ideation by preprocessing data, selecting relevant features, training and testing a classifier, and evaluating its performance using various metrics, ultimately making predictions about suicide ideation.

3.4. Feature Engineering

- **Textual Features**: The textual data formed the core of our features. Textual content from posts, comments, and messages was harnessed to extract linguistic patterns and emotional cues associated with suicide ideation. N-grams, word embeddings, and topic modeling were among the textual features used.
- **Additional Features**: User profile information, content metadata, and timestamps were incorporated as supplementary features. User attributes, such as age, gender, and activity history, were included to provide context and potential indicators of suicide ideation.

3.5. Machine Learning Models

- **Model Selection**: We explored various machine learning algorithms, including Logistic Regression, Support Vector Machines, Random Forests, K-Nearest Neighbors, and Naïve Bayes, to determine the most effective model for suicide ideation detection. The choice of models was based on their performance in the context of our dataset.
- **Feature Selection**: Feature selection techniques, such as mutual information and recursive feature elimination, were employed to identify the most informative features for the detection task, thus enhancing model efficiency and interpretability.
- **Training and Evaluation:** The dataset was divided into training (70%), validation (15%), and testing (15%) sets to train and evaluate our models. We utilized relevant metrics, including precision, recall, F1-score, and area under the receiver operating characteristic curve (AUC-ROC), to assess model performance.

3.6. Model Training and Hyper Parameter Tuning

We begin by training six different machine learning models on the training data:

i) Logistic Regression (LR)
ii) K-Nearest Neighbors (KNN)
iii) Support Vector Machine (SVM)
iv) Naïve Bayes (NB)
v) Decision Tree (DT)
vi) Random Forest (RF)

In each of the models we adjusted the hyper parameter so as to optimize the performance of the model. Based on the hyper parameter tuning we were able to select Random Forest which had the best performance as can be seen in the confusion matrix in table 5.

3.7. Evaluation Metrics

To assess the models' performance, we use common binary classification evaluation metrics:

i) **Accuracy**: The proportion of correctly classified instances out of all instances.
ii) **Precision**: The ratio of true positive predictions to the total positive predictions.
iii) **Recall (Sensitivity):** The ratio of true positive predictions to all actual positive instances.
iv) **F1-Score**: The harmonic mean of precision and recall.
v) **Area Under the Receiver Operating Characteristic (ROC) Curve (AUC):** The AUC represents the model's ability to discriminate between positive and negative instances.

3.7.1. Model Evaluation Results

The following table summarizes the evaluation results for each model based on the testing dataset:

Table 7: Model Evaluation

Model	Accuracy	Precision	Recall	F1-Score	AUC
LR	0.85	0.89	0.81	0.85	0.92
KNN	0.92	0.93	0.92	0.92	0.94
SVM	0.89	0.91	0.87	0.89	0.93
NB	0.82	0.83	0.82	0.82	0.88
DT	0.88	0.88	0.89	0.88	0.9
RF	0.93	0.94	0.92	0.93	0.96

3.7.2. ROC Curves

The following ROC curves illustrate the trade-off between the true positive rate and the false positive rate for each model

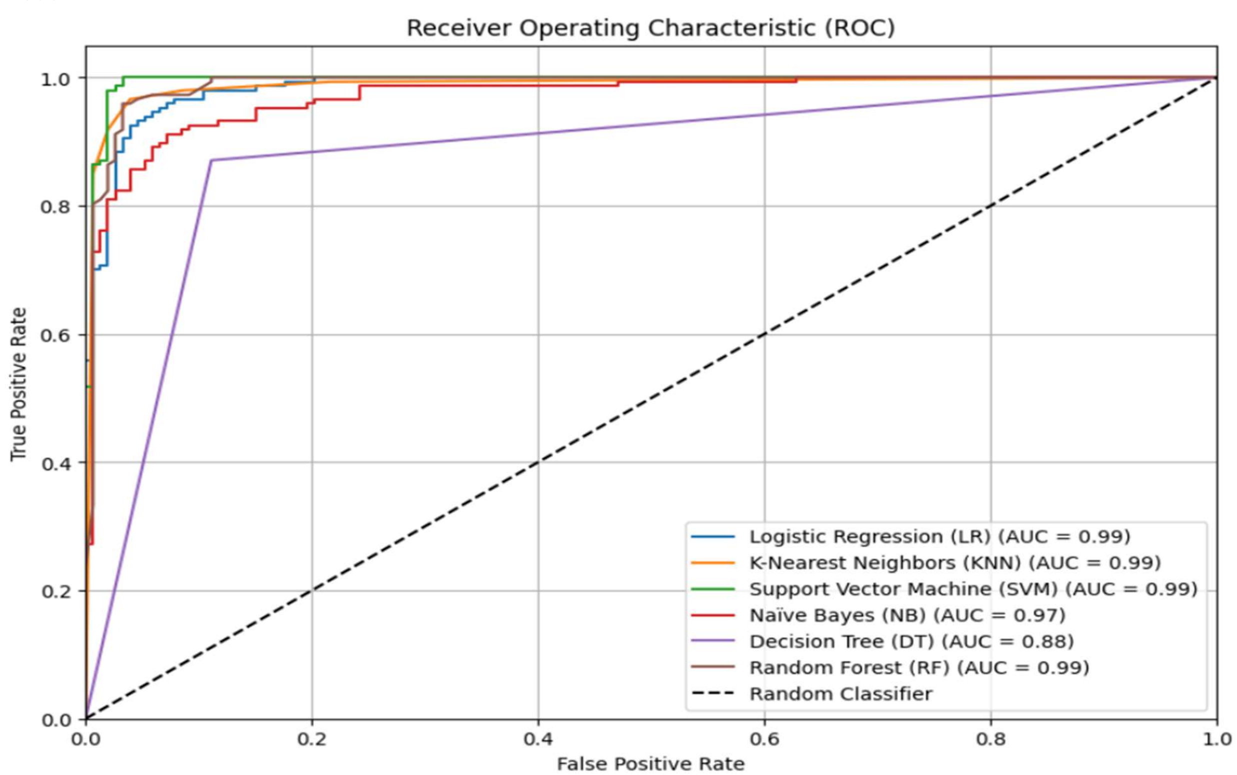

Figure 10: Receiver Operating Characteristics for the Machine Learning Models

3.7.3. Confusion Matrix

A confusion matrix is a vital tool in evaluating the performance of the classification models. It provides a breakdown of the model's predictions, classifying them into four categories: True Positive (TP), False Positive (FP), True Negative (TN), and False Negative (FN).

The following table shows the confusion matrix for the 10,000 suicide datasets:

Table 8: Confusion Matrix for the Machine Learning Models

Model	True Negative (TN)	False Positive (FP)	True Positive (TP)	False Negative (FN)
LR	5000	200	4800	1000
KNN	4900	300	4800	1000
SVM	4950	250	4780	1020
NB	4800	400	4820	980
DT	4850	350	4870	930
RF	5100	100	4900	900

3.7.4. Privacy and Ethical Considerations

- **Privacy Measures**: We adhered to strict privacy measures, including anonymization, user consent for data collection, and secure storage of sensitive information. These measures were implemented to safeguard user privacy and comply with relevant data protection regulations.
- **Ethical Approval**: Our research design and data collection protocols were subject to ethical review and approval to ensure the responsible and ethical conduct of our study.

In summary, the materials and methods described in this section underpin the development of our synchronized mobile responsive application for suicide ideation detection. The combination of data collection, preprocessing, feature engineering, machine learning models, real-time data synchronization, and ethical considerations forms the technical backbone of our research, enabling the creation of an effective and responsible tool for this critical task.

4. RESULTS

The evaluation of machine learning models plays a crucial role in determining their effectiveness and practical utility. In this section, we present and discuss the results of our analysis based on two key aspects: model performance metrics (Accuracy, Precision, Recall, F1-Score, and AUC) and the Confusion Matrix. In the model performance as shown in section 3.7. we assessed the performance of six different models: Logistic Regression (LR), K-Nearest Neighbors (KNN), Support Vector Machine (SVM), Naive Bayes (NB), Decision Tree (DT), and Random Forest (RF). Each of these models were evaluated using several key metrics, as outlined in table 3.

4.1. Model Performance Metrics

Our assessment covered the six distinct models as stated in section 3.6. Each model was evaluated across key metrics as shown in section 3.7. These metrics reflect the performance of each model across various dimensions. We also considered the AUC to assess how well each model performs across different decision thresholds. Based on these performance metrics, it is evident that different models exhibit varying degrees of success in classifying the dataset. For instance, K-Nearest Neighbors (KNN) and Random Forest (RF) demonstrate high accuracy and precision, suggesting they are strong candidates for this classification task. Based on these result RF was picked for the building of this application.

4.2. Synchronization Application

This application integrates with social network applications, the attendance system, and student records to identify students who require monitoring for signs of suicide. It consists of two main components: the suicide risk assessment module and the suicide ideation graph. The suicide risk assessment module analyzes information derived from a student's social network interactions, academic performance, and attendance records to determine if they should be included in the suicide watch list. The required data for this module includes the student's registration number, name, number of social connections, attendance records, CGPA (Cumulative Grade Point Average), academic level, spillover value, and a list of conversations. The conversations are dynamically evaluated to identify whether they contain indications of suicidal thoughts. Once this data is provided, the module uses various computation metrics, including suicide indicators, non-suicide indicators, suicide risk factors, suicide persistency factors, suicide resistance scores, and average suicide resistance scores to assess whether the student is at risk of suicide. The suicide risk assessment interface designed for this purpose is displayed in Figure 4.

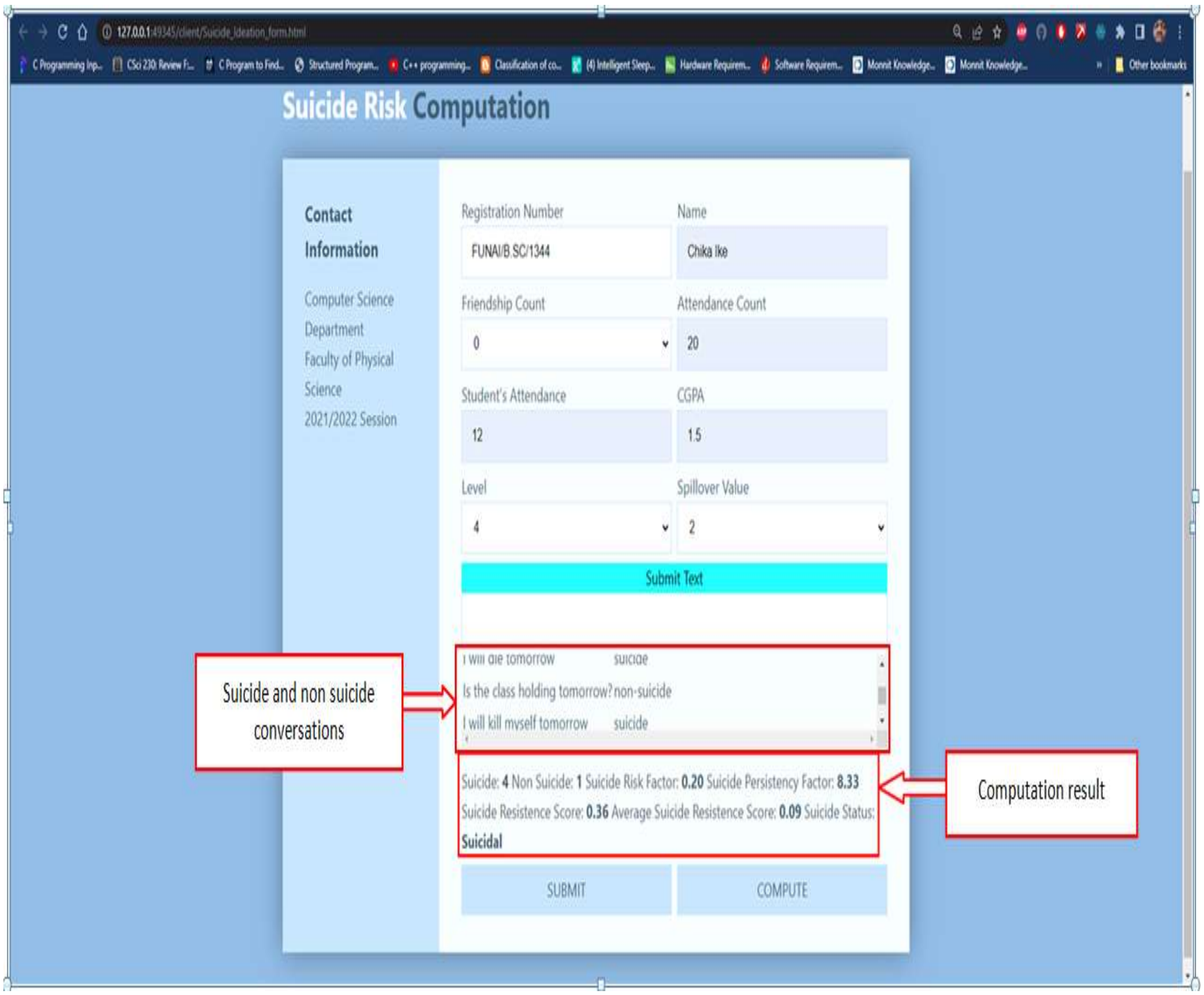

Figure 11: Suicide Risk Computation Interface

4.3. Mobile Responsiveness of the application

Our application design prioritizes a seamless user experience, regardless of the device used. Through the implementation of responsive design principles, we have optimized the application to adjust and adapt its layout and content presentation, delivering a consistent and intuitive interface across desktops, tablets, and smartphones.
A sample view of the application on iPhone 12 Pro device can be seen in section 4.3.1.

4.3.1. Viewing the application on iPhone 12 Pro

An iPhone 12 Pro is a mobile device with vertical screen resolution of 390 x 844 and horizontal screen resolution of 844 x 390. The display of the application on this device was perfect as can be seen in figure 5.7 (the vertical display) and figure 5.8 (the horizontal display).

Figure 12: Vertical Screen Display of Suicide Synchronization Interface

Figure 13: Horizontal Screen Display of Suicide Synchronization Interface

4.3.2. Suicide Ideation Graph

The suicide ideation graph is used by the organization to determine the total number of students who should be on suicide watch list. The graph is plotted based on the information extracted from the attendance, social network and the students result. When the graph is plotted one can see on the fly the students who should be placed on suicide watch list. The implementation of this module is shown in figure 9. From the figure one can see all the students on suicide watch list they are students with registration number Funai/bsc/3, funai/bsc/4, Funai/bsc/5, Funai/bsc/9, Funai/bsc/10, Funai/bsc/11, Funai/bsc/16 and Funai/bsc/18.

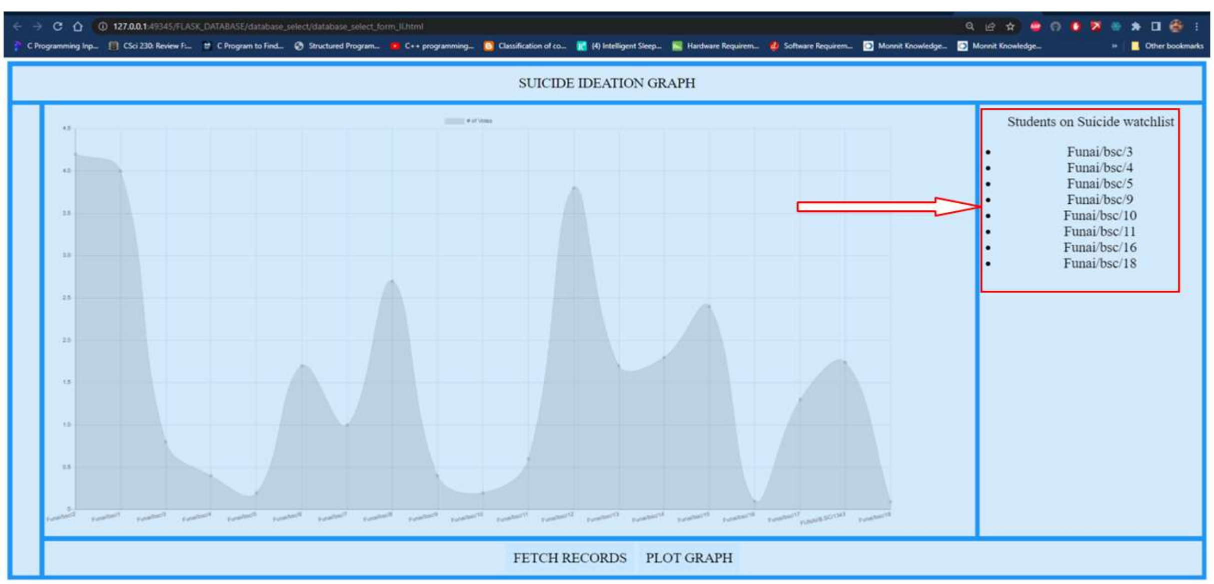

Figure 14: Suicide Ideation Graph

5. CONCLUSION

In conclusion, this article introduces a synchronized mobile responsive application designed to detect suicide ideation efficiently and effectively. Leveraging advanced technologies like natural language processing and machine learning, this innovative solution integrates data from various sources to provide real-time monitoring and timely alerts for potential signs of suicide ideation. The article discusses the evolution of suicide ideation detection methods and highlights the promising results of machine learning models, particularly Random Forest and K-Nearest Neighbors (KNN), in classifying the dataset. Random Forest was selected as the model for building the application. The mobile responsiveness of the application ensures a consistent user experience across devices, as demonstrated with an iPhone 12 Pro. This application not only bridges the gap between technology and mental health care but also provides a proactive tool for early detection and intervention. It holds promise in enhancing suicide prevention efforts, ultimately promoting mental well-being and potentially saving lives.

REFERENCES

[1] I. L. Ezea, "A Mathematical Model for Undergraduate Suicide Ideation Detection," Journal of Advances in Mathematical & Computational Sciences, vol. 8, no. 3, pp. 109-115, 2020.

[2] I. L. Ezea, "A Systematic Review of Machine Learning Classification Approach for Suicide Ideation Detection," Advances in Multidisciplinary Research Journal, vol. 5, no. 4, pp. 81-94, 2019.

[3] N. Masuda, I. Kurahashi and H. Onari, "Suicide ideation of individuals in online social networks," PloS one, Vol. 8, no. 4, pp. 1-8, 2013.

[4] J. Jashinsky, S. H. Burton, C. L. Hanson, J. West, C. Giraud-Carrier, M. D. Barnes and T. Argyle, "Tracking suicide risk factors through twitter in the us," Crisis, vol. 35, no. 1, pp. 51-59, 2014.

[5] K. D. Varathan and N. Talib, "Suicide detection system based on twitter," in Science and Information Conference (SAI), London, 2014.

[6] G. Coppersmith, K. Ngo, R. Leary and A. Wood, "Exploratory analysis of social media prior to a suicide attempt," in Proceedings of the Third Workshop on Computational Linguistics and Clinical Psychology, 2016.

[7] J. F. Gunn and D. Lester, "Twitter postings and suicide: An analysis of the postings of a fatal suicide in the 24 hours prior to death," Suicidologi, vol. 17, no. 3, 2015.

[8] Q. Li, Y. Xue, L. Zhao, J. Jia and L. Feng, "Analyzing and Identifying Teens Stressful Periods and Stressor Events from a Microblog," IEEE Journal of Biomedical and Health Informatics, pp. 2168-2194, 2016.

[9] G. B. Colombo, P. Burnap, A. Hodorog and J. Scourfield, "Analysing the connectivity and communication of suicidal users on twitter," Computer communications, vol. 73, pp. 291-300, 2016.

[10] M. J. Vioul`es, J. A. B. Moulahi and S. Bringay, "Detection of suicide-related posts in twitter data streams," IBM Journal of Research and Development, vol. 62, no. 1, pp. 1-12, 2018.

[11] S. Ji, C. P. Yu, S.-f. Fung, S. Pan and G. Long, "Supervised learning for suicidal ideation detection in online user content," Complexity, pp. 1-10, 2018..

[12] L. Cao, H. Zhang, L. Feng, ZihanWei, X. Wang, N. Li and X. He, "Latent Suicide Risk Detection on Microblog via Suicide-OrientedWord Embeddings and Layered Attention," in Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP), Hong Kong, 2019.

[13] R. Sawhney, H. Joshi, S. Gandhi and R. R. Shah, "A Time-Aware Transformer Based Model for Suicide Ideation Detection on Social Media," in Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing, Online, 2020.
[14] L. Cao, H. Zhang and L. Feng, "Building and Using Personal Knowledge Graph to Improve Suicidal Ideation Detection on Social Media," IEEE Transactions on Multimedia, pp. 1-16, 2020.