



International Journal of Innovative Research in Science Engineering and Technology (IJIRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)



Impact Factor: 8.699

Volume 14, Issue 11, November 2025

A Hybrid Predictive Model for Identifying Suicidal Ideation from User-Generated Text

N. Sowmmeeya, Prof. S. Raghul, K.N.Sivakumar

M.E Student, Department of Computer Science and Engineering, M.P. Nachimuthu M.Jaganathan Engineering College,
Chennimalai, Erode, Tamilnadu, India

Assistant Professor, Department of Computer Engineering, M.P.Nachimuthu M.Jaganathan Engineering College,
Chennimalai, Erode, Tamilnadu, India

Associate Professor, Department of CSE, M.P.Nachimuthu M.Jaganathan Engineering College, Chennimalai, Erode,
Tamilnadu, India

ABSTRACT: Mental health challenges, including suicidal tendencies, pose serious risks to individuals' well-being and require timely intervention. Accurate risk prediction and early detection are therefore crucial for providing effective mental health support. This study presents a comparative framework of multiple machine learning models for suicidal risk prediction. Logistic Regression and Random Forest classifiers were trained on a large dataset of textual data collected from mental health forums and social media, while a Simple RNN model was trained to capture sequential and contextual patterns in the text. For all models, the dataset was pre-processed using text cleaning, lemmatization, and vectorization or sequence encoding. Model performance was evaluated using classification metrics, demonstrating strong predictive capabilities. The resulting system is integrated web application, offering an interactive chatbot platform where users can safely communicate their thoughts and receive real-time supportive responses. The hybrid model further combines predictions from all individual models using majority voting, enhancing accuracy and reliability. These results highlight the effectiveness of combining machine learning and deep learning models for mental health risk detection, providing a practical tool for timely intervention, awareness, and emotional support.

I. INTRODUCTION

Suicide is a leading cause of premature mortality, potentially preventable [80]. The World Health Organization (WHO) reports that more than 700,000 individuals die by suicide each year, that is, one death every 40 seconds [113]. Also, around 77% of suicides in the world occur in low- and middle-income countries. This phenomenon impacts individuals across all regions and age groups worldwide [8], [29], [81]. Suicide is influenced by a complex interplay of factors such as media coverage, economic conditions, biological influences, psychology issues, environmental factors, history context, and social dynamics [83], [90], [108]. Individuals at high risk for suicide often struggle with mental health disorders, such as bipolar disorder, anxiety, and depression [18], [73], [88]. The persistent experience of extreme hopelessness, helplessness, worthlessness, or feelings of defeat and confinement gradually drives individuals toward suicidal behaviour [86], [90], [108].

Suicidal behaviour includes a spectrum of actions ranging from Suicidal Ideation (SI) and planning to attempts and completed suicides [13]. SI, which precedes actual suicide, encompasses thoughts, desires, and preoccupations related to death and self-harm. This includes planning, writing farewell notes, and even expressing these feelings through social media [42], [74]. SI is associated with a variety of mental health disorders, such as alcohol use disorder, panic symptoms, anxiety, and sleep disturbances, as well as situational stressors [13], [42]. Among older adults, cultural acceptance of suicide, religiosity, and intergenerational cohabitation were predictors of SI [13]. Social media has increasingly become a central platform for people to express themselves and interact with others [58] For vulnerable groups, such as those struggling with mental health issues or substance abuse problems, these platforms may be used to discuss and acquire information about suicide [93].

Effective suicide prevention requires the active involvement of Mental Health Professionals (MHPs), such as psychologists and psychiatrists [62]. Traditionally, SI analysis involves psychological evaluations, which include direct

patient-MHP interactions and suicide risk identification based on self-reports and face-to-face interviews [38], [56], [98]. However, these approaches depend heavily on in-person interactions, requiring extensive training, ongoing supervision, workflow adjustments, and considerable time and effort, which limits their scalability [16]. Digital interventions based on Artificial Intelligence (AI) are developed to improve mental health support. These novel approaches include clinical decision support systems like *Boamente* [31], [32], which was proposed to detect and monitor SI remotely through patients' mobile devices (e.g., smartphones and tablets). The *Boamente* system passively collects non-clinical texts written in Brazilian Portuguese (PT-BR) from patients through a customized virtual keyboard application.

These texts are transmitted to a web platform, where an AI model classifies them as indicative of negative or positive for SI (i.e., a binary classification task). The web application then presents these classifications on dashboards with graphical representations, allowing MHPs to monitor remotely patients' SI.

II. LITERATURE SURVEY

Background: In 2016, 44,965 people in the United States died by suicide. It is common to see people with suicidal ideation seek help or leave suicide notes on social media before attempting suicide. Many prefer to express their feelings with longer passages on forums such as Reddit and blogs. Because these expressive posts follow regular language patterns, potential suicide attempts can be prevented by detecting suicidal posts as they are written. **Objective:** This study aims to build a classifier that differentiates suicidal and non-suicidal forum posts via text mining methods applied on post titles and bodies. **Methods:** A total of 508,398 Reddit posts longer than 100 characters and posted between 2008 and 2016 on Suicide Watch, Depression, Anxiety, and Shower Thoughts subreddits were downloaded from the publicly available Reddit dataset. Of these, 10,785 posts were randomly selected and 785 were manually annotated as suicidal or non-suicidal. Features were extracted using term frequency-inverse document frequency, linguistic inquiry and word count, and sentiment analysis on post titles and bodies. Logistic regression, random forest, and support vector machine (SVM) classification algorithms were applied on resulting corpus and prediction performance is evaluated. **Results:** The logistic regression and SVM classifiers correctly identified suicidality of posts with 80% to 92% accuracy and F1 score, respectively, depending on different data compositions closely followed by random forest, compared to baseline ZeroR algorithm achieving 50% accuracy and 66% F1 score.

Introduction: Around a million people are reported to die by suicide every year, and due to the stigma associated with the nature of the death, this figure is usually assumed to be an underestimate. Suicide may be prevented if prompt intervention is taken to mitigate risk. Machine learning and artificial intelligence-based modelling, such as natural language processing (NLP) and other text analytics approaches, has the potential to become a major technique for the detection, diagnosis, and treatment of people who are suffering from mental health issues. The primary aims of this research are to determine whether NLP techniques have been utilised in the field of suicide prevention, and if so, were they effective? What were their limitations?

Methods: PubMed, EMBASE, MEDLINE, PsycINFO, and Global Health databases were searched for studies that reported use of NLP for suicide ideation or self-harm. Thematic analysis was used to synthesise and analyse the included studies. Findings were reported using the Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA) statement, and the Mixed Methods Appraisal Tool (MMAT) was used in assessing paper quality. **Result:** The preliminary search of five databases generated 387 results. Removal of duplicates resulted in 158 potentially suitable studies. Twenty papers were finally included in this review. **Discussion:** Studies show that combining structured and unstructured data in NLP data modelling yielded more accurate results than utilizing either alone. Also, to reduce suicides, people with mental problems must be continuously and passively monitored. Further, NLP and other machine learning/artificial intelligence technologies can be used to address health inequities and electronic health records provide valuable data for creating suicide risk tools. Finally, Online, social media, and smartphone applications can be leverage in detecting people with suicide ideation. **Conclusion:** The use of artificial intelligence and machine learning opens new avenues for considerably guiding risk prediction and advancing suicide prevention frameworks. The review's analysis of the included research revealed that the use of NLP may result in low-cost and effective alternatives to existing resource-intensive methods of suicide prevention. To summarise, there is substantial evidence that NLP is useful in identifying people who have suicide ideation.

Artificial intelligence (AI) can transform health care practices with its increasing ability to translate the uncertainty and complexity in data into actionable—though imperfect—clinical decisions or suggestions. In the evolving relationship between humans and AI, trust is the one mechanism that shapes clinicians' use and adoption of AI. Trust is a psychological mechanism to deal with the uncertainty between what is known and unknown. Several research studies have highlighted the need for improving AI-based systems and enhancing their capabilities to help clinicians. However, assessing the magnitude and impact of human trust on AI technology demands substantial attention. Will a clinician trust an AI-based system? What are the factors that influence human trust in AI? Can trust in AI be optimized to improve decision-making processes? In this paper, we focus on clinicians as the primary users of AI systems in health care and present factors shaping trust between clinicians and AI.

A large number of individuals are suffering from suicidal ideation in the world. As the most popular platform for self-expression, emotion release, and personal interaction, individuals may exhibit a number of symptoms of suicidal ideation on social media. Nevertheless, challenges from both data and knowledge aspects remain as obstacles, constraining the social media-based detection performance. Data implicitness and sparsity make it difficult to discover the inner true intentions of individuals based on their posts. Inspired by psychological studies, we build and unify a high-level suicide-oriented knowledge graph with deep neural networks for suicidal ideation detection on social media. We further design a two-layered attention mechanism to explicitly reason and establish key risk factors to individual's suicidal ideation. The performance study on microblog and Reddit shows that: 1) with the constructed personal knowledge graph, the social media-based suicidal ideation detection can achieve over 93% accuracy; and 2) among the six categories of personal factors, post, personality, and experience are the top-3 key indicators. Under these categories, posted text, stress level, stress duration, posted image, and ruminant thinking contribute to one's suicidal ideation detection. Suicide remains a critical global health issue, with over 800,000 deaths annually and profound impacts on individuals and their communities. The increasing role of social media platforms in expressing personal struggles has highlighted the potential of these platforms for early detection of suicidal tendencies.

Advances in machine learning and natural language processing (NLP) have revolutionized this detection process, allowing for more accurate identification of suicidal ideation in social media posts. This review examines the integration of machine learning algorithms in suicide detection, including the use of multi-layered neural networks for analysing textual, audio, and visual data. It explores various machine learning models and their performance in predicting suicide attempts, with a focus on their accuracy, precision, and recall. The review also discusses applications such as real-time monitoring systems and emergency response mechanisms, emphasizing how these technologies can enhance early intervention and support. By surveying current research and identifying challenges and future directions, this review aims to provide a comprehensive understanding of how machine learning can significantly contribute to suicide prevention efforts through effective social media monitoring.

III. PROBLEM STATEMENT

Mental health issues, including stress, anxiety, and suicidal tendencies, are increasingly prevalent worldwide, impacting individuals' well-being, productivity, and social interactions. Early detection and intervention are critical to prevent escalation and provide timely support. Traditional approaches, such as self-report questionnaires and clinical consultations, are often limited by accessibility, stigma, and the delay between symptom onset and professional intervention. Automated systems for mental wellness support face challenges in accurately interpreting user-generated textual data, as the language can be nuanced, context-dependent, and emotionally complex. Simple keyword-based or rule-based approaches often fail to capture the subtle cues indicative of mental distress or suicidal tendencies. Additionally, most existing tools provide either basic assessments or generic responses, limiting their practical effectiveness for personalized support.

To address these challenges, this project integrates multiple machine learning models—Logistic Regression, Random Forest, and a Simple RNN neural network—to analyse user messages and predict mental health risk levels. The system classifies input into “Non-Suicidal” or “Suicidal” categories with associated confidence scores, enhancing interpretability and reliability. By combining predictive analytics with a interactive web application, this project provides a scalable, real-time, and user-friendly platform for mental wellness support, helping users receive timely guidance and professional referral when needed.

IV. PROPOSED METHOD

The proposed system introduces an intelligent hybrid framework that leverages both machine learning (ML) and deep learning (DL) techniques to accurately predict suicidal risk based on users' textual inputs. Unlike traditional or rule-based systems, this approach captures not only word-level meaning but also contextual, emotional, and sequential patterns in language, enabling deeper insight into users' mental health states. The system integrates three powerful models — XGBoost, Naive Bayes, Random Forest, and Simple Recurrent Neural Network (RNN) — to form a hybrid majority voting model. By combining these models, the hybrid system ensures improved accuracy, robustness, and reliability in detecting suicidal tendencies.

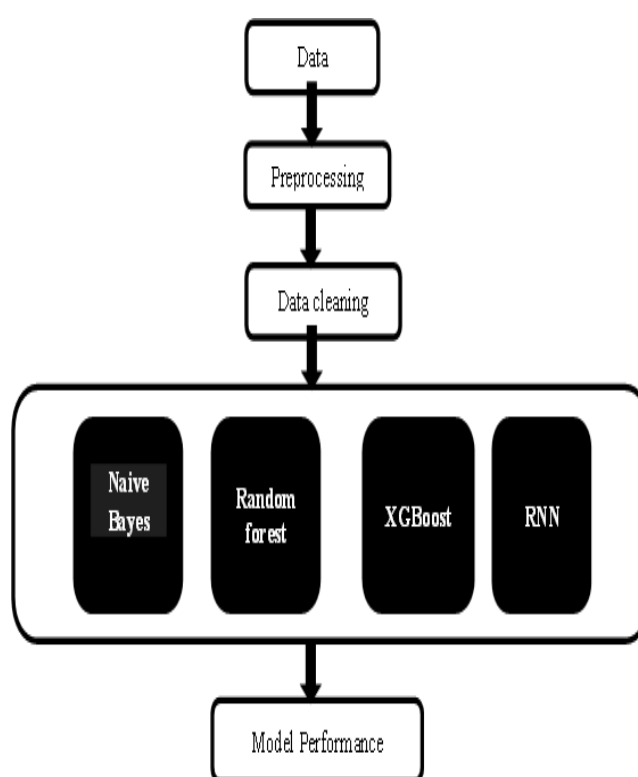


Figure. 1 SYSTEM ARCHITECTURE

The experimental results demonstrate the effectiveness of various machine learning and deep learning models for suicidal risk prediction using text-based data. Four models—Naive Bayes, Random Forest, XGBoost, and Simple RNN—were trained and evaluated on the preprocessed dataset comprising 80,000 samples. Among the traditional machine learning models, XGBoost achieved the highest accuracy of 90.94%, followed by Naive Bayes with 90.26% and Random Forest with 90.06%, indicating that ensemble and probabilistic methods perform effectively in identifying suicidal intent from textual content. However, the Simple RNN model outperformed all other models with an accuracy of 92.42%, showing its superior ability to capture sequential dependencies and contextual meaning in natural language.

The classification reports further revealed balanced performance across both “suicide” and “non-suicide” classes, with precision, recall, and F1-scores all above 0.90, indicating minimal bias and strong generalization. The confusion matrix visualizations confirmed that the models were consistent in correctly classifying most samples, with few false positives and false negatives. The comparative accuracy chart highlights that deep learning models leveraging sequential information can significantly improve prediction accuracy in text-based mental health analysis. Overall, the findings suggest that hybrid frameworks combining traditional ML algorithms with deep learning architectures can serve as a robust and reliable approach for early suicidal risk detection.

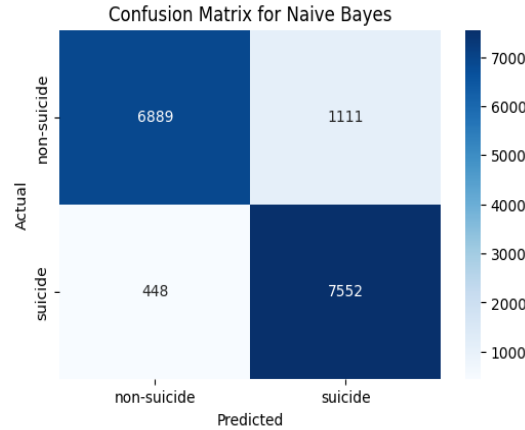


Figure.2 Naïve Bayes Confusion Matrix

Figure 2 illustrate the confusion matrix representing the performance of the Naïve Bayes in predicting accuracy level. TF-IDF is a technique used to convert text data into numerical feature vectors. It assigns a numerical weight to each word in a document based on its frequency within that document and its rarity across the entire collection of documents (Inverse Document Frequency). This process highlights words that are important within a specific document but are not overly common across all documents, making them good indicators for distinguishing between document categories.

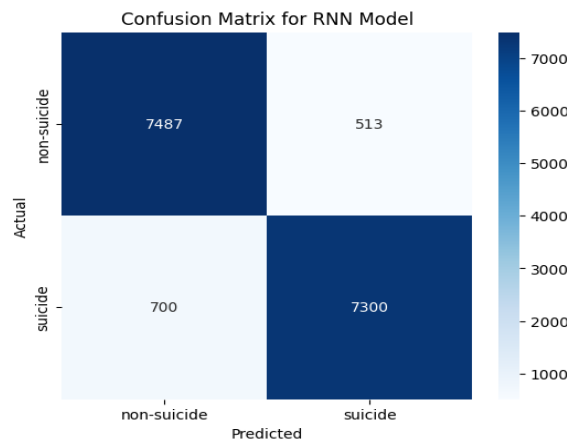


Figure 3 RNN Model Confusion Matrix

Figure 3 illustrate the confusion matrix representing the performance of the RNN Model in predicting accuracy level. An RNN, or Recurrent Neural Network, is a type of artificial neural network in machine learning that is designed to process sequential data, like text, speech, or time series. Its key feature is an internal "memory" that allows it to retain information from previous inputs, enabling it to make predictions or classifications based on the context of the sequence. This makes RNNs useful for tasks such as language translation, speech recognition, and sentiment analysis.

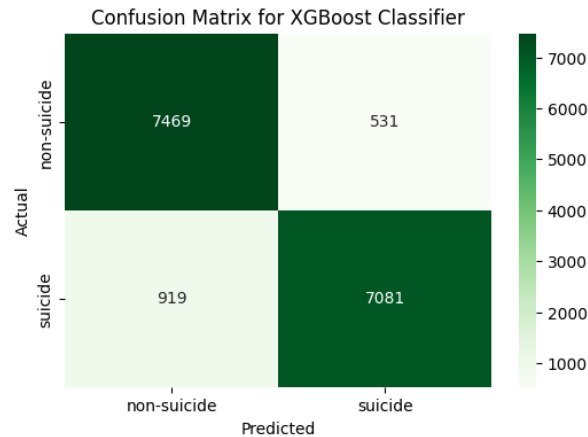


Figure. 4 XG Boost Confusion Matrix

Figure 4 illustrate the confusion matrix representing the performance of the XG Boost Model in predicting accuracy level. XGBoost, which stands for eXtreme Gradient Boosting, is an open-source machine learning library and an advanced implementation of the gradient boosting algorithm. It is widely used for its efficiency, flexibility, and performance in various machine learning tasks, including classification and regression.

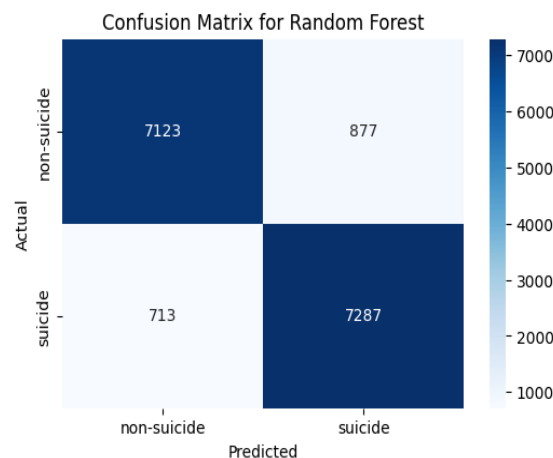


Figure.5 Random Forest Confusion Matrix

Random Forest Algorithm:

Random Forest (RF) is an ensemble learning method that combines multiple decision trees to improve classification accuracy and reduce overfitting. Each tree in the forest is trained on a random subset of the training data and features, and the final prediction is determined through majority voting.

$$\hat{y} = \text{mode}\{h_1(x), h_2(x), h_3(x), \dots, h_n(x)\}$$

$\hat{y}$ is the final predicted class,

$h_1(x)$ Represents the prediction made by the i^{th} decision tree, and

n denotes the total number of trees in the forest.

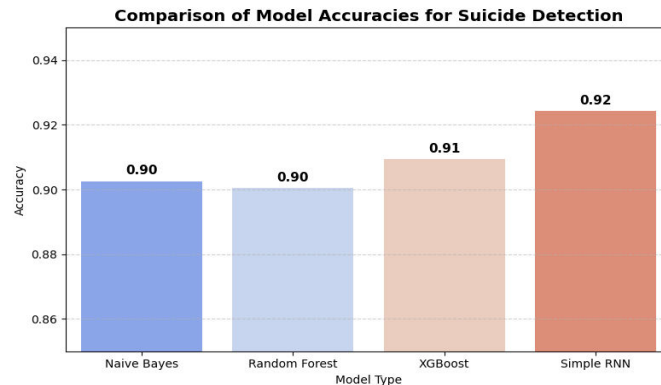


Figure. 6 Comparison of Accuracy

Figure 6 illustrates the comparison of model accuracies between the Naïve Bayes, XGBoost, RNN and Random Forest algorithms used for mental health classification. The Naïve Bayes model achieved an accuracy of 90.26%, demonstrating strong performance in predicting risk categories. However, the Random Forest model outperformed it with an impressive accuracy of 90.06%, Optimized DL Model in simple RNN achieved an accuracy of 92.42%, indicating a nearly perfect classification rate and XGBoost Model achieved an accuracy of 90.94%, demonstrating strong and powerful performance in predicting risk categories. The significant improvement in the Random Forest model's performance can be attributed to its ensemble learning mechanism, which combines multiple decision trees to reduce variance and enhance generalization.

This comparison highlights that the RNN model provides a more robust and reliable approach for predicting mental health risk levels compared to other three models.

V. CONCLUSION

The proposed system demonstrates the potential of machine learning models in supporting mental health and suicide prevention efforts. By analysing user input messages, the trained classifier effectively distinguishes between suicidal and non-suicidal expressions with high accuracy. The integration of a chatbot web application provides users with a simple, interactive platform to express their feelings and receive immediate feedback. The system is further strengthened by an email alert mechanism, which automatically notifies registered contacts when suicidal tendencies are detected. These alerts include messages, helpline numbers, and motivational quotes supportive, ensuring timely intervention and emotional support.

Additionally, the platform supports confidence scores, message history, and interpretability features, making it more transparent and user-friendly. Overall, this project highlights how AI-powered tools, combined with thoughtful design, can provide real-time mental wellness support, encourage early awareness, and potentially save lives by guiding users toward professional help.

REFERENCES

- [1] A. Abid, A. Abdalla, A. Abid, D. Khan, A. Alfozan, and J. Zou, "Gradio: Hassle-free sharing and testing of ML models in the wild," 2019, arXiv:1906.02569.
- [2] A. Adadi and M. Berrada, "Peeking inside the black-box: A survey on explainable artificial intelligence (XAI)," IEEE Access, vol. 6, pp. 52138–52160, 2018.
- [3] A. Adadi and M. Berrada, "Explainable AI for healthcare: From black box to interpretable models," in Embedded Systems and Artificial Intelligence, V. Bhateja, S. C. Satapathy, and H. Satori, Eds., Singapore: Springer, 2020, pp. 327–337.
- [4] A. Adjekum, A. Blasimme, and E. Vayena, "Elements of trust in digital health systems: Scoping review," J. Med. Internet Res., vol. 20, no. 12, Dec. 2018, Art. no. e11254.

- [5] A. E. Aladağ, S. Muderrisoglu, N. B. Akbas, O. Zahmacioglu, and H. O. Bingol, "Detecting suicidal ideation on forums: Proof-of-concept study," *J. Med. Internet Res.*, vol. 20, no. 6, Jun. 2018, Art. no. e215.
- [6] S. Ali, T. Abuhmed, S. El-Sappagh, K. Muhammad, J. M. Alonso- Moral, R. Confalonieri, R. Guidotti, J. Del Ser, N. Díaz-Rodríguez, and F. Herrera, "Explainable artificial intelligence (XAI): What we know and what is left to attain trustworthy artificial intelligence," *Inf. Fusion*, vol. 99, Nov. 2023, Art. no. 101805.
- [7] A. A. Vanderbilt, N. J. Pastis, J. Mayglothling, and D. Franzen, "A review of the literature: Direct and video laryngoscopy with simulation as educational intervention," *Adv. Med. Educ. Pract.*, vol. 5, pp. 15–23, Jan. 2014.
- [8] R. Alonzo, J. Hussain, S. Stranges, and K. K. Anderson, "Interplay between social media use, sleep quality, and mental health in youth: A systematic review," *Sleep Med. Rev.*, vol. 56, Apr. 2021, Art. no. 101414.
- [9] American Psychiatric Association, *DSM-5: Diagnostic and Statistical Manual of Mental Disorders*, A. Editora, Ed., Washington, DC, USA: American Psychiatric Association, 2014.
- [10] A. Arowosegbe and T. Oyelade, "Application of natural language processing (NLP) in detecting and preventing suicide ideation: A systematic review," *Int. J. Environ. Res. Public Health*, vol. 20, no. 2, p. 1514, Jan. 2023.
- [11] O. Asan, A. E. Bayrak, and A. Choudhury, "Artificial intelligence and human trust in healthcare: Focus on clinicians," *J. Med. Internet Res.*, vol. 22, no. 6, Jun. 2020, Art. no. e15154. [12] L. Balcombe and D. De Leo, "Digital mental health challenges and the horizon ahead for solutions," *JMIR Mental Health*, vol. 8, no. 3, Mar. 2021, Art. no. e26811.
- [13] S. Booniam, T. Wongpakaran, P. Lerttrakarnnon, S. Jiraniramai, P. Kuntawong, and N. Wongpakaran, "Predictors of passive and active suicidal ideation and suicide attempt among older people: A study in tertiary care settings in Thailand," *Neuropsychiatric Disease Treatment*, vol. 16, pp. 3135–3144, Dec. 2020.
- [14] A. Bourla, F. Ferreri, L. Ogorzelec, C.-S. Peretti, C. Guinchard, and S. Mouchabac, "Psychiatrists' attitudes toward disruptive new technologies: Mixed-methods study," *JMIR Mental Health*, vol. 5, no. 4, Dec. 2018, Art. no. e10240.
- [15] A. Bourla, S. Mouchabac, L. Ogorzelec, C. Guinchard, and F. Ferreri, "Are student nurses ready for new technologies in mental health? Mixedmethods study," *Nurse Educ. Today*, vol. 84, Jan. 2020, Art. no. 104240.



INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

 9940 572 462  6381 907 438  ijirset@gmail.com



www.ijirset.com

Scan to save the contact details