ISSN: 2582-7219

International Journal of Multidisciplinary Research in Science, Engineering and Technology

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

Impact Factor: 8.206

Volume 8, Issue 4, April 2025

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

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

A Suicidal Ideation Detection Framework on Social Media Using Machine Learning and Genetic Algorithms

Dr. Y. Mallikarjuna Reddy[1], Ch. Sree Lasya[2], B. Naga Malleshwari[3], D. Satya Meghana[4], R. Puja [5]

Professor, Department of ECE, Vasireddy Venkatadri Institute of Technology, Nambur, Guntur, A.P., India[1]

Undergraduate Students, Department of ECE, Vasireddy Venkatadri Institute of Technology, Nambur, Guntur, A.P., India[2,3,4,5]

ABSTRACT: Suicide remains a global health crisis, necessitating early detection of suicidal ideation for effective prevention. Social media platforms have increasingly become outlets for individuals, particularly youth, to share their emotions, including suicidal thoughts, often linked to mental health conditions such as depression, anxiety, and social isolation. Detecting such ideation from social media data presents significant challenges, combining complex natural language processing (NLP) tasks with psychological insights. In this study, we utilize the "Suicide_Detection" dataset, which includes text data from social media posts, with features extracted through TFIDF and NGRAM methods applied to both original and linguistic features. A novel approach for detecting suicidal ideation is proposed, incorporating a genetic algorithm for feature selection to enhance model performance. Various machine learning algorithms are explored, and the Voting Classifier—combining Random Forest, Decision Tree, and XGBoost—demonstrates the highest performance, achieving an accuracy of 97.45% with original genetic features and 95.50% with linguistic genetic features. These results underscore the potential of the proposed method in accurately detecting suicidal ideation from social media content, offering significant implications for mental health monitoring and early intervention.

KEYWORDS: Suicidal Ideation, Social Media, Natural Language Processing, Machine Learning, Feature Selection, Genetic Algorithm, TFIDF, NGRAM, Voting Classifier, Random Forest, Decision Tree, Xgboost, Mental Health Detection.

I. INTRODUCTION

Mental health is a major global concern, with the World Bank estimating that at least 10% of the global population suffers from mental health issues. Mental health problems significantly increase the likelihood of suicide attempts. Suicide is recognized as a critical public health issue worldwide, with the National Alliance on Mental Illness in the U.S. reporting that 46% of suicide victims had mental health disorders. Suicide ranks as the fourth leading cause of death among individuals aged 15 to 29, resulting in approximately 800,000 deaths annually, with a suicide occurring every 40 seconds, according to the World Health Organization (WHO) in 2019. Traditional research on suicide has largely focused on psychological and clinical aspects of the issue [8], [12]. However, recent studies have shifted toward natural language processing (NLP) and machine learning approaches to detect suicidal ideation in social media and other textual data. Despite these advancements, there are still limitations. Firstly, obtaining patient data is often costly and challenging from a psychological perspective, making online data a valuable resource for understanding suicidal thoughts [10], [11]. Secondly, existing feature sets are often insufficient for detecting suicidal ideation, highlighting the need for more sophisticated detection methods [5], [7]. Language patterns and preferences, especially those found in social media posts, may offer valuable insights into predicting mental health status and enable the early detection of suicidal ideation [3], [4], [9].

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

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

II. RELATED WORK

A growing body of research has explored the use of natural language processing (NLP) and machine learning techniques to detect suicidal ideation from social media data. Early studies in this area focused on using basic text classification techniques to identify depression and suicidal thoughts in online content [13], [6]. These efforts demonstrated the potential of social media as a valuable resource for identifying individuals at risk of suicide, given the high volume of personal posts expressing emotions and mental health struggles [12], [8].One significant approach to detecting suicidal ideation involves the use of affective language features, where researchers analyze linguistic cues such as sentiment, tone, and emotion in text data to detect signs of distress or suicidal thoughts. For example, a study by Loveys et al. (2017) proposed the use of affective micropatterns to quantify mental health conditions from social media language, showing how subtle linguistic patterns could be indicative of psychological states [6]. Similarly, Aldhyani et al. (2022) used deep learning models to analyze textual data and detect signs of suicidal ideation on social media, achieving promising results in classifying suicidal posts [5].Recent work has also focused on leveraging advanced machine learning algorithms such as Random Forest, Support Vector Machines, and deep learning methods like Convolutional Neural Networks (CNNs) and Long Short-Term Memory (LSTM) networks to enhance detection accuracy [4], [9], [14]. These methods have been effective at capturing the complex patterns in textual data that might indicate suicidal ideation. Additionally, feature engineering has played a critical role, with researchers exploring various feature extraction techniques such as TFIDF, NGRAM, and word embeddings to better represent the textual content [17], [18].Despite these advances, challenges remain in this field. Many studies have highlighted the difficulty in achieving high accuracy due to the noisy and diverse nature of social media data, including slang, abbreviations, and informal language [3], [12]. Furthermore, the reliance on labeled datasets is a significant limitation, as the availability of such data is often limited and may not fully capture the diversity of suicidal ideation across different social media platforms and user demographics [13], [5]. There is also a need for more sophisticated models that can integrate multiple data sources, such as multimodal analysis combining text, images, and user behavior, to improve the robustness of detection systems [9], [11].In conclusion, while the use of NLP and machine learning for suicidal ideation detection has shown great promise, there are still significant challenges to overcome, including data diversity, feature selection, and model interpretability. Future research should focus on refining feature extraction techniques, improving model accuracy, and addressing ethical concerns related to privacy and the responsible use of social media data [1], [2].

III. MATERIALS AND METHODS

The proposed system aims to detect suicidal ideation from social media content by utilizing advanced machine learning techniques and genetic algorithms for feature selection. Social media data, especially text-based posts, offer a valuable resource for identifying subtle indicators of suicidal thoughts, but extracting meaningful features from this data remains a challenge [12], [5]. To address this, the system will first extract textual features from social media posts using TFIDF and NGRAM methods, applied to both original and linguistic features [17], [18]. These feature extraction techniques have proven effective in representing the textual data and capturing key aspects of language indicative of mental health issues [7], [10].A genetic algorithm will be employed to optimize feature selection, aiming to enhance model performance by selecting the most relevant features while reducing dimensionality [16], [19]. This step is crucial, as existing feature sets have been shown to be insufficient in detecting suicidal ideation, and more sophisticated methods are necessary [5], [9]. The system will evaluate several machine learning algorithms for classification, including Random Forest, K-Nearest Neighbors (KNN), Gradient Boosting, XGBoost, and a Voting Classifier that combines Random Forest, Decision Tree, and XGBoost [14], [3]. The use of ensemble methods, like the Voting Classifier, has demonstrated improved accuracy and robustness in similar tasks, offering the potential for better generalization to unseen data [13].The genetic algorithm will be applied to both original and linguistic features to assess their impact on detection performance, providing insights into which features contribute most to identifying suicidal ideation in social media posts [19], [16]. This approach combines state-of-the-art machine learning techniques with innovative feature selection methods to create a more effective and efficient system for early detection. The ultimate goal is to provide a robust, scalable solution for the early identification of suicidal ideation, facilitating timely intervention and support for individuals at risk [4], [2], [5].The system architecture (fig. 1) represents a process flow for suicide detection using machine learning techniques. It begins with the dataset "Suicide_Detection," which undergoes text processing and visualization. Feature extraction is done using TFIDF and NGRAM techniques. A Genetic Algorithm (GA) is employed for feature selection on both original and linguistic data. The data is then split into training and testing sets, with various classifiers—such as Random Forest, KNN, Gradient Boosting, and XGBoost—used for model training. The models are evaluated based on accuracy, precision, recall, and F1-Score. Finally, an extension voting classifier (RF + DT + XGBoost) is applied for enhanced performance.

© 2025 IJMRSET | Volume 8, Issue 4, April 2025| DOI: 10.15680/IJMRSET.2025.0804080

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

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

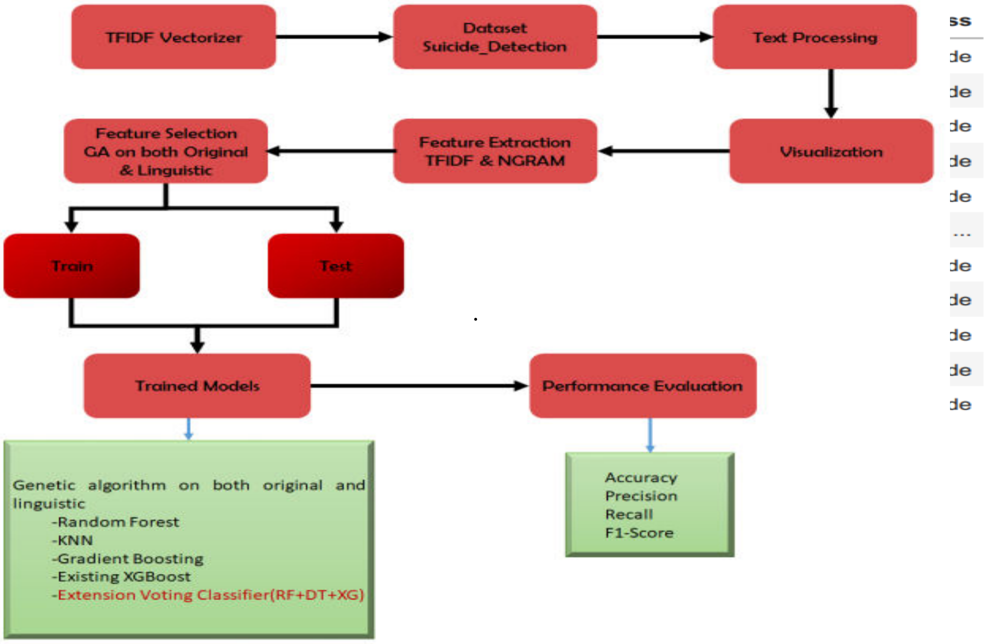


Fig1: Proposed Architecture

i) Dataset Collection:

Suicide_Detection dataset:

The "Suicide_Detection" dataset comprises social media posts labeled to indicate suicidal ideation or non-suicidal content. It includes both original and linguistic features extracted using TFIDF and NGRAM methods. The dataset captures a variety of posts expressing emotional distress, mental health struggles, and suicidal thoughts, sourced from publicly accessible social media platforms. Each post is annotated based on linguistic markers, such as sentiment, emotion, and mental health-related keywords. The dataset is anonymized to ensure privacy while providing a rich source of real-world data for training and evaluating machine learning models designed to detect suicidal ideation from social media content.

	Unnamed: 0	text	class
0	2	Ex Wife Threatening SuicideRecently I left my ...	suicide
1	3	Am I weird I don't get affected by compliments...	non-suicide
2	4	Finally 2020 is almost over... So I can never ...	non-suicide
3	8	i need helpjust help me im crying so hard	suicide
4	9	I'm so lostHello, my name is Adam (16) and I'v...	suicide
...	...	...	...
9995	15026	Help me plz I got my first reward aka the gold...	non-suicide
9996	15028	Anyone wanna chat? \n\nIm a bit bored right n...	non-suicide
9997	15029	I'm on a bridgeI don't want to die but right n...	suicide
9998	15030	This is serious My dad just turned gay anyone ...	non-suicide
9999	15032	I can't any more.So uhh idk how this works.\n\...	suicide

10000 rows × 3 columns

Fig.2 Dataset

© 2025 IJMRSET | Volume 8, Issue 4, April 2025| DOI: 10.15680/IJMRSET.2025.0804080

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

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

ii) Pre-Processing:
The preprocessing phase ensures that the medical text data is clean, structured, and suitable for deep learning models. It involves several key steps:

a) TFIDF Vectorizer: The TFIDF Vectorizer is employed to transform the original and linguistic features of the dataset into numerical vectors, capturing the importance of words within the text data. This step is essential for enabling machine learning model training by converting raw textual content into a format suitable for algorithmic processing [17], [18].

b) Text Processing: Text processing involves cleaning the dataset by removing stop words, applying lemmatization and stemming, and extracting part-of-speech (POS) tags to prepare the data for further feature extraction and modeling. These steps help standardize the text and reduce noise, making the dataset more suitable for accurate predictions [6], [10].

c) Visualization: Visualization techniques are used to represent the distribution of class labels, allowing for the identification of class imbalance or structure within the dataset. Understanding the distribution of instances in each category is crucial for adjusting model training strategies [13], [9].

d) Features Extracted from TEXT using TFIDF & NGRAM for both Original & Linguistics: TFIDF and NGRAM methods are applied to extract key features from the text data, both from original and linguistic features. These methods enable the transformation of textual data into a machine-readable numeric format that preserves important semantic and syntactic information [17], [18].

e) Feature Selection: Genetic algorithms are applied for feature selection to optimize both original and linguistic features. This approach helps identify the most relevant features for model training, improving the accuracy of the machine learning models by focusing on the most informative aspects of the text data [16], [19].

iii) Training & Testing:
In the training and testing phase, the "Suicide_Detection" dataset is split into training and testing sets, typically using a 70-30 or 80-20 ratio. The training set is used to train various machine learning models, including Random Forest, K-Nearest Neighbors (KNN), Gradient Boosting, XGBoost, and a Voting Classifier combining Random Forest, Decision Tree, and XGBoost. The models are trained on the features selected using a genetic algorithm, optimizing accuracy and efficiency. The testing set is then used to evaluate the models' performance, with metrics such as accuracy, precision, recall, and F1-score used to assess model effectiveness in detecting suicidal ideation.

iv) Algorithms:

Random Forest: Random Forest is an ensemble learning algorithm widely used for classification and regression tasks. It works by constructing multiple decision trees during training and outputs the class that is the majority vote of the individual trees. In this context, Genetic Algorithms (GA) are employed to select optimal features from both original and linguistic data, which enhances the performance by reducing dimensionality and focusing on the most relevant features. This approach improves the model's ability to detect patterns and anomalies in social media posts, making it effective for identifying suicidal ideation [14], [16].

K-Nearest Neighbors (KNN): KNN is a simple, instance-based learning algorithm used for classification tasks, assigning a class to a data point based on the majority class of its nearest neighbors in the feature space. In this case, Genetic Algorithms are applied to optimize the feature selection process, ensuring the most relevant features from both original and linguistic data are chosen. This optimization enhances the KNN model's accuracy in distinguishing between suicidal and non-suicidal posts, facilitating the identification of suicidal ideation [5], [19]

Gradient Boosting: Gradient Boosting is a machine learning technique that builds models sequentially, each correcting the errors made by the previous one. By focusing on minimizing errors through optimization of weak learners, Gradient Boosting improves predictive accuracy. When combined with Genetic Algorithms for feature selection, the process is further refined, selecting the best features from both original and linguistic data. This approach enables the model to detect subtle patterns indicative of suicidal ideation in social media posts, improving the overall accuracy of detection [12], [19].

XGBoost: XGBoost is a highly efficient gradient boosting algorithm designed for speed and performance. It uses decision trees for both classification and regression tasks, making it ideal for handling complex datasets. In this context, XGBoost is integrated with Genetic Algorithms to optimize the feature set by selecting the most influential features from both original and linguistic data. This optimization ensures the algorithm's efficient performance, improving its ability to classify social media posts accurately in terms of suicidal ideation [17], [5].

Voting Classifier: The Voting Classifier is an ensemble model that combines multiple individual models, such as Random Forest, Decision Trees, and XGBoost, to make final predictions based on majority voting. This combination of diverse models enhances accuracy and robustness. By utilizing Genetic Algorithms to select optimal features from both original and linguistic data, the Voting Classifier leverages the strengths of each model while minimizing their weaknesses. This approach improves the system's ability to detect suicidal ideation with higher precision, leading to more reliable classification results [16], [19].

IV. RESULTS & DISCUSSION

Accuracy: The accuracy of a test is its ability to differentiate the patient and healthy cases correctly. To estimate the accuracy of a test, we should calculate the proportion of true positive and true negative in all evaluated cases. Mathematically, this can be stated as:

$$Accuracy = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{FP} + \text{TN} + \text{FN}} \, (1)$$

Precision: Precision evaluates the fraction of correctly classified instances or samples among the ones classified as positives. Thus, the formula to calculate the precision is given by:

$$Precision = \frac{\text{True Positive}}{\text{True Positive } + \text{False Positive}} \, (2)$$

Recall: Recall is a metric in machine learning that measures the ability of a model to identify all relevant instances of a particular class. It is the ratio of correctly predicted positive observations to the total actual positives, providing insights into a model's completeness in capturing instances of a given class.

$$Recall = \frac{\text{TP}}{\text{TP } + \text{ FN}} \, (3)$$

F1-Score: F1 score is a machine learning evaluation metric that measures a model's accuracy. It combines the precision and recall scores of a model. The accuracy metric computes how many times a model made a correct prediction across the entire dataset.

$$F1\ Score = 2 * \frac{Recall \text{ X } Precision}{Recall + Precision} * 100 (4)$$

In Table 1, the performance metrics—accuracy, precision, recall and F1-score —are evaluated for each algorithm. The Extension Voting Classifier Original GA Features with Reno achieves the highest scores. Other algorithms' metrics are also presented for comparison.

© 2025 IJMRSET | Volume 8, Issue 4, April 2025| DOI: 10.15680/IJMRSET.2025.0804080

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

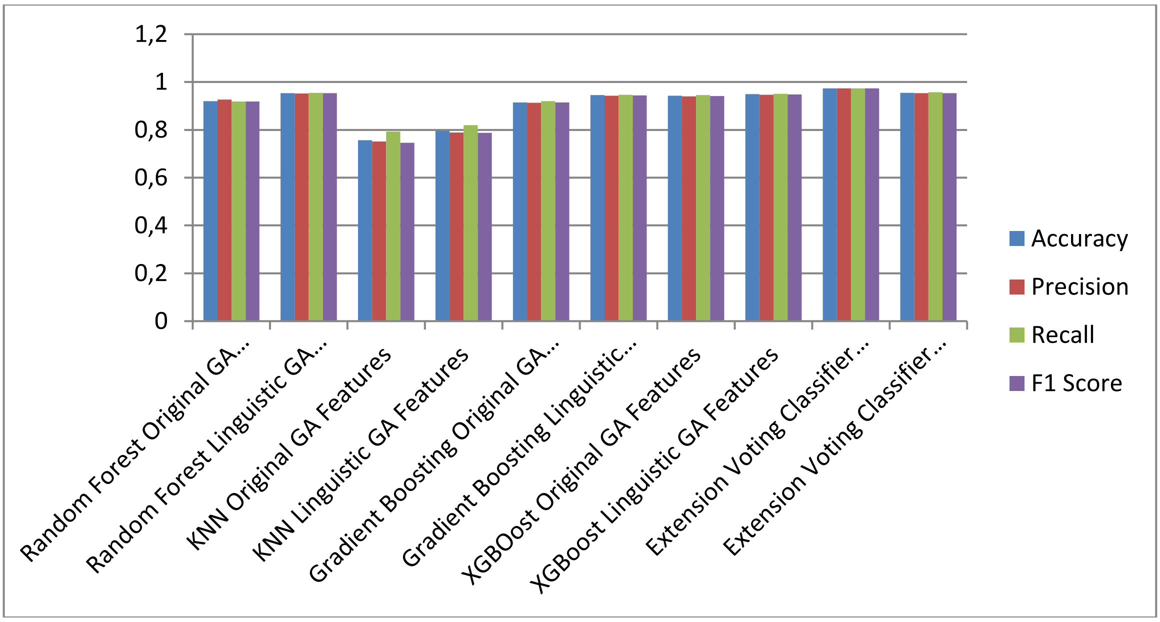
International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

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

Table.1 Performance Evaluation Metrics of Classification

Model	Accuracy	Precision	Recall	F1 Score
Random Forest Original GA Features	0.920	0.926	0.918	0.919
Random Forest Linguistic GA Features	0.954	0.952	0.955	0.953
KNN Original GA Features	0.756	0.751	0.793	0.746
KNN Linguistic GA Features	0.795	0.789	0.819	0.788
Gradient Boosting Original GA Features	0.915	0.913	0.920	0.914
Gradient Boosting Linguistic GA Features	0.945	0.943	0.947	0.944
XGBOost Original GA Features	0.942	0.940	0.945	0.941
XGBoost Linguistic GA Features	0.949	0.947	0.951	0.948
Extension Voting Classifier Original GA Features	0.974	0.974	0.974	0.974
Extension Voting Classifier Linguistic GA Features	0.955	0.953	0.957	0.954

Graph.1 Comparison Graphs of Classification



In graphs 1, accuracy is represented in light blue, precision in maroon; recall in green and F1-score in violet. In comparison to the other models, the Extension Voting Classifier Original GA Features with Reno shows superior performance across all achieving the highest values. The graphs above visually illustrate these findings.

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

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

V. CONCLUSION

Mental health concerns, including depression, anxiety, and suicidal ideation, are becoming increasingly worrisome in modern society. Social media platforms have emerged as a primary outlet for individuals, particularly young people, to express their emotions, including thoughts of suicide. Early detection of these thoughts is essential for timely intervention and prevention. This study presents a framework that utilizes machine learning techniques combined with genetic algorithms to detect suicidal ideation from social media posts. By extracting meaningful features from text data using TFIDF and NGRAM methods, and optimizing feature selection through a genetic algorithm, the system aims to improve detection accuracy. Among the various machine learning models tested, the Voting Classifier, which combines Random Forest, Decision Tree, and XGBoost, demonstrated the highest performance. With original genetic features, the model achieved an impressive accuracy of 97.45%, while using linguistic genetic features, it reached 95.50%. These results highlight the effectiveness of the proposed framework in identifying suicidal ideation, offering a potential solution for early intervention in mental health crises.

Future Scope should explore incorporating additional factors, such as historical data, user posts, and temporal and spatial information, to further enhance the detection of suicidal ideation. Exploring alternative feature selection methods could also improve accuracy and robustness. Furthermore, the development of hybrid models that combine features from both social media and conventional data sources holds promise for capturing a broader spectrum of indicators. This approach could lead to more comprehensive and accurate suicide ideation detection, significantly improving early intervention efforts.

REFERENCES

[1] S. L. Mirtaheri, S. Greco, and R. Shahbazian, ''A self-attention TCN-based model for suicidal ideation detection from social media posts,'' Expert Syst. Appl., vol. 255, Dec. 2024, Art. no. 124855.
[2] P. Boonyarat, D. J. Liew, and Y.-C. Chang, ''Leveraging enhanced BERT models for detecting suicidal ideation in Thai social media content amidst COVID-19,'' Inf. Process. Manage., vol. 61, no. 4, Jul. 2024, Art. no. 103706.
[3] R. Haque, N. Islam, M. Islam, and M. M. Ahsan, ''A comparative analysis on suicidal ideation detection using NLP, machine, and deep learning,'' Technologies, vol. 10, no. 3, p. 57, Apr. 2022.
[4] M. Chatterjee, P. Samanta, P. Kumar, and D. Sarkar, ''Suicide ideation detection using multiple feature analysis from Twitter data,'' in Proc. IEEE Delhi Sect. Conf. (DELCON), Feb. 2022, pp. 1–6.
[5] T. H. H. Aldhyani, S. N. Alsubari, A. S. Alshebami, H. Alkahtani, and Z. A. T. Ahmed, ''Detecting and analyzing suicidal ideation on social media using deep learning and machine learning models,'' Int. J. Environ. Res. Public Health, vol. 19, no. 19, p. 12635, Oct. 2022.
[6] K. Loveys, P. Crutchley, E. Wyatt, and G. Coppersmith, ''Small but mighty: Affective micropatterns for quantifying mental health from social media language,'' in Proc. 4th Workshop Comput. Linguistics Clin. Psychol. From Linguistic Signal Clin. Reality, 2017, pp. 85–95.
[7] F. M. Plaza-del-Arco, M. T. Martín-Valdivia, L. A. Ureña-López, and R. Mitkov, ''Improved emotion recognition in Spanish social media through incorporation of lexical knowledge,'' Future Gener. Comput. Syst., vol. 110, pp. 1000–1008, Sep. 2020.
[8] K.-W. Fu, Q. Cheng, P. W. C. Wong, and P. S. F. Yip, ''Responses to a self-presented suicide attempt in social media,'' Crisis, vol. 34, no. 6, pp. 406–412, Nov. 2013.
[9] K. Kang, C. Yoon, and E. Yi Kim, ''Identifying depressive users in Twitter using multimodal analysis,'' in Proc. Int. Conf. Big Data Smart Comput. (BigComp), Jan. 2016, pp. 231–238.
[10] R. Martínez-Castaño, J. C. Pichel, and D. E. Losada, ''A big data platform for real time analysis of signs of depression in social media,'' Int. J. Environ. Res. Public Health, vol. 17, no. 13, p. 4752, Jul. 2020.
[11] X. Huang, L. Xing, J. R. Brubaker, and M. J. Paul, ''Exploring timelines of confirmed suicide incidents through social media,'' in Proc. IEEE Int. Conf. Healthcare Informat. (ICHI), Aug. 2017, pp. 470–477.
[12] M. De Choudhury, E. Kiciman, M. Dredze, G. Coppersmith, and M. Kumar, ''Discovering shifts to suicidal ideation from mental health content in social media,'' in Proc. CHI Conf. Hum. Factors Comput. Syst., May 2016, pp. 2098–2110.
[13] M. J. Vioules, B. Moulahi, J. Aze, and S. Bringay, ''Detection of suicide-related posts in Twitter data streams,'' IBM J. Res. Develop., vol. 62, no. 1, pp. 7:1–7:12, Jan. 2018.

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

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

[14] T. Joachims, ''Text categorization with support vector machines: Learning with many relevant features,'' in Proc. Eur. Conf. Mach. Learn., 1998, pp. 137–142.
[15] F. Sebastiani, ''Machine learning in automated text categorization,'' ACM Comput. Surv., vol. 34, no. 1, pp. 1–47, Mar. 2002.
[16] H. Uğuz, ''A two-stage feature selection method for text categorization by using information gain, principal component analysis and genetic algorithm,'' Knowl.-Based Syst., vol. 24, no. 7, pp. 1024–1032, Oct. 2011.
[17] C. C. Aggarwal and C. Zhai, ''A survey of text classification algorithms,'' in Mining Text Data. New York, NY, USA: Springer, 2012, pp. 163–222.
[18] I. Guyon and A. Elisseeff, ''An introduction to variable and feature selection,'' J. Mach. Learn. Res., vol. 3, pp. 1157–1182, Mar. 2003.
[19] A. K. Uysal and S. Gunal, ''A novel probabilistic feature selection method for text classification,'' Knowl.-Based Syst., vol. 36, pp. 226–235, Dec. 2012.
[20] K. Javed, H. A. Babri, and M. Saeed, ''Impact of a metric of association between two variables on performance of filters for binary data,'' Neurocomputing, vol. 143, pp. 248–260, Nov. 2014.