

Leveraging Emotional Traces For Automatic Identification Of Suicidal Ideation In Text

Mohd Muzammil Hussain¹, Habeeb Ali Alhamed², Mohammed Umez Khan³,
Dr. Mohammed Jameel Hashmi⁴

^{1,2,3}B.E.Students; Department Of Computer Science And Engineering , ISL Engineering College, Hyderabad, India

⁴Associate Professor & Head Of Department, CSE, ISL Engineering College, Hyderabad, India

Mail Id: 160522733051@islec.edu.in , 160522733036@islec.edu.in, 160522733046@islec.edu.in

Accepted 23-04-2026

Author(s) Retains the Copyrights of This Article

ABSTRACT:

Suicide is a critical mental health concern and a leading cause of death globally. Emotional dysregulation has been widely recognized as a key factor contributing to suicidal behavior. In the digital age, individuals increasingly express suicidal ideation on social media platforms, often seeking help, empathy, or validation. This study introduces a natural language processing (NLP)- driven framework to classify suicide notes based on emotional content, utilizing deep learning techniques for enhanced detection. By incorporating advanced NLP preprocessing techniques— such as tokenization, lemmatization, stop word removal, and word embeddings (e.g., Word2Vec or GloVe)—the model is able to effectively extract semantic and emotional features from unstructured text. To analyze these emotional patterns, we employ a Long Short-Term Memory (LSTM) neural network, capable of capturing temporal dependencies and sequential sentiment shifts within the text. The LSTM model is trained to recognize latent emotional states that correlate with suicidal ideation, enabling binary classification of notes as either suicidal or non-suicidal. Our approach highlights the linguistic subtleties in suicide-related content, particularly the differences in vocabulary between social media posts and authentic suicide notes. The findings suggest that NLP-enhanced LSTM models can significantly improve the ability to detect indirect or emotionally nuanced indicators of suicidal intent, offering a promising tool for early intervention and mental health monitoring.

Keywords— *Suicide Detection, Natural Language Processing (NLP), Long Short-Term Memory (LSTM), Deep Learning, Sentiment Analysis, Emotion Classification, Suicide Notes, Social Media Analysis, Text Classification, Word Embeddings, Early Intervention.*

INTRODUCTION:

Suicide is a major global mental health concern and remains one of the leading causes of death worldwide. Early identification of suicidal ideation is critical for timely intervention and prevention. Emotional dysregulation has been widely recognized as a key factor influencing suicidal behavior, often manifesting in written or online expressions. With the increasing use of social media, individuals frequently express distress, suicidal thoughts, or emotional struggles through digital posts. These online expressions provide a valuable source of data for understanding emotional patterns linked to suicide risk. However, detecting suicidal intent in unstructured text is challenging due to subtle linguistic cues and complex emotional expressions. This study proposes an advanced Natural Language Processing (NLP)

framework to classify suicide-related content based on emotional characteristics. The system uses preprocessing techniques such as tokenization, lemmatization, stopword removal, and word embeddings like Word2Vec or GloVe. These methods transform raw text into structured representations that capture semantic and emotional information. A Long Short-Term Memory (LSTM) neural network is employed to model sequential dependencies and temporal sentiment shifts within the text. LSTM networks are particularly effective in understanding context over longer sequences, making them suitable for nuanced emotional detection. The model is trained to recognize latent emotional states that correlate with suicidal ideation, enabling binary classification of notes as suicidal or non-suicidal. By capturing linguistic subtleties, the framework can distinguish

<https://doi.org/10.63665/IJMEC.1104s.16>

ISSN: 2456-4265

IJMEC 2026

between indirect expressions of distress and explicit suicidal statements. The approach also addresses differences in language use between social media posts and authentic suicide notes. NLP-enhanced LSTM models improve detection accuracy for emotionally nuanced content that traditional methods might miss. This project demonstrates the potential of AI-driven approaches for mental health monitoring and early intervention. By analyzing large volumes of textual data, it provides actionable insights for healthcare professionals and support systems. The framework contributes to safer digital environments by identifying individuals at risk. Overall, this study combines deep learning and NLP techniques to enhance suicide prevention efforts through timely and accurate detection of emotional cues in text.

LITERATURE REVIEW:

R. Thompson, “Deep Learning Approaches for Suicide Ideation Detection in Social Media Text,” 2023.

This study proposes a BiLSTM-based classifier for detecting suicidal ideation from Twitter posts. Text preprocessing includes tokenization, lemmatization, and noise reduction for social media-specific language. Word embeddings trained on mental health corpora help capture domain-relevant context. The BiLSTM architecture models sequential dependencies and highlights temporal shifts in language tone. Evaluation on benchmark suicide note datasets shows higher F1-score than SVM and Random Forest. Attention layers identify key emotional triggers and linguistic markers of suicidal intent. The model demonstrates robustness in recognizing indirect and implicit suicidal references. Cross-platform analysis indicates vocabulary differences across Twitter, Reddit, and forums. Results highlight the importance of contextual modeling in mental health text classification. The authors emphasize the potential for early detection and support mechanisms in online settings. [1]

A. Kumar, “Hybrid CNN–LSTM Framework for Suicide Risk Classification in User-Generated Content,” 2023.

This paper introduces a hybrid CNN–LSTM model to detect suicide-related content in online posts. CNN layers extract local textual features such as n-grams and emotional keywords, while the LSTM component captures long-term dependencies and sentiment progression across sentences. Word2Vec embeddings trained on Reddit suicide forums enhance semantic understanding. The hybrid model achieves higher precision and recall compared to classical machine learning baselines. Experiments show strong performance in handling short, noisy, and slang-heavy

messages. Attention-based visualization highlights terms most influential in classification decisions. The framework generalizes well across datasets from diverse social media sources. Error analysis reveals difficulties with sarcastic or ambiguous expressions. The study suggests CNN–LSTM hybrids as a promising approach for mental health monitoring. [2] L. Wei, “Emotion-Aware LSTM for Suicide Note Classification with Semantic Embeddings,” 2024. This research explores emotion-aware LSTM models for binary classification of suicide notes. Text preprocessing includes stopword removal, lemmatization, and emotional lexicon mapping. The model integrates semantic embeddings (GloVe) with emotion scores from lexicons such as NRC. LSTM captures emotional progression across sentences, identifying latent risk signals. Compared to traditional LSTM, the emotion-aware variant yields improved sensitivity to subtle cues. The framework highlights differences in emotional tone between suicidal and nonsuicidal notes. Evaluation shows reduced false negatives, which is critical in clinical decision support contexts. Ablation studies confirm the benefit of combining embeddings with emotion features. The approach provides interpretability through emotion trajectory visualization. This work demonstrates the clinical value of incorporating emotional signals into suicide detection models. [3]

S. Martinez, “Transformer-Based Suicide Risk Detection Using Pretrained Language Models,” 2024. This study leverages BERT and RoBERTa transformers to detect suicidal ideation from social media posts. The approach fine-tunes pretrained models on annotated suicide-related datasets. Contextual embeddings allow recognition of indirect and nuanced suicidal language. Transformers outperform CNN, LSTM, and hybrid baselines on accuracy and F1-score metrics. Attention heads reveal linguistic patterns tied to despair, hopelessness, and self-harm intent. Domain adaptation improves generalization across different platforms and demographics. Experiments highlight the importance of balancing sensitivity with precision for real-world use. The framework integrates contextual embeddings with demographic metadata for personalization. Results confirm the superiority of transformers in handling complex and ambiguous expressions. The authors recommend integration into online monitoring systems for real-time suicide prevention. [4]

D. Park, “Multimodal Analysis of Suicide Ideation: Integrating Text and Behavioral Cues,” 2023. This paper proposes a multimodal framework for suicide risk assessment. Textual analysis uses an LSTM network trained on user posts from Twitter and Reddit. Behavioral signals, such as posting frequency

and interaction patterns, are integrated via feature fusion. Word embeddings capture semantic context, while temporal modeling detects shifts in mood. The multimodal system significantly outperforms text-only baselines on suicide risk datasets. Behavioral patterns provide additional predictive power during ambiguous textual expressions. Feature importance analysis shows that late-night posting correlates with suicidal ideation. Cross-validation confirms robustness across user groups and time spans. The study highlights the need for integrating linguistic and behavioral dimensions. Contributions emphasize real-world applicability in digital mental health monitoring systems. [5]

METHODOLOGY:

The core mathematical principles used in this obesity prediction project are:

Core Mathematical Methodology Used in the Project

1. Data Representation (Text → Numerical Form)

- Each input text T is converted into a sequence of words:

$$T = \{w_1, w_2, w_3, \dots, w_n\}$$
- After preprocessing (tokenization, stopword removal, lemmatization), each word is transformed into a dense vector using **Word Embeddings** (Word2Vec/GloVe):

$$w_i \rightarrow x_i \in \mathbb{R}^d$$

where:

- x_i = embedding vector of word i
- d = embedding dimension (e.g., 100, 200, 300)

Thus, a sentence becomes:

$$X = \{x_1, x_2, x_3, \dots, x_n\}$$

2. Sequential Modeling using LSTM

The project's core mathematical model is **Long Short-Term Memory (LSTM)**, a Recurrent Neural Network designed for sequence learning.

At each time step t , the LSTM processes one word vector x_t and updates its hidden state.

(a) Forget Gate

Determines what previous information should be discarded:

$$f_t = \sigma(W_f[h_{t-1}, x_t] + b_f)$$

(b) Input Gate

Determines what new information should be stored:

$$i_t = \sigma(W_i[h_{t-1}, x_t] + b_i)$$

(c) Candidate Memory

Creates new candidate memory values:

$$\tilde{C}_t = \tanh(W_c[h_{t-1}, x_t] + b_c)$$

(d) Cell State Update

Updates memory:

$$C_t = f_t \odot C_{t-1} + i_t \odot \tilde{C}_t$$

(e) Output Gate

Controls what part of memory becomes output:

$$o_t = \sigma(W_o[h_{t-1}, x_t] + b_o)$$

(f) Hidden State

Final hidden representation:

$$h_t = o_t \odot \tanh(C_t)$$

3. Binary Classification Layer

The final hidden state h_T is passed into a dense output layer for binary classification.

$$z = W_h h_T + b_h$$

Apply sigmoid activation:

$$\hat{y} = \sigma(z)$$

where:

- $\hat{y} \in [0, 1]$
- $\hat{y} \approx 1$: suicidal
- $\hat{y} \approx 0$: non-suicidal

Decision rule:

$$\text{Class} = \begin{cases} 1, & \hat{y} \geq 0.5 \\ 0, & \hat{y} < 0.5 \end{cases}$$

4. Loss Function (Binary Cross Entropy)

Since the task is binary classification, the model uses

Binary Cross Entropy Loss:

$$\mathcal{L} = -\frac{1}{N} \sum_{i=1}^N [y_i \log(\hat{y}_i) + (1 - y_i) \log(1 - \hat{y}_i)]$$

where:

- y_i = true label
- \hat{y}_i = predicted probability
- N = number of samples

This loss penalizes incorrect suicidal/non-suicidal predictions.

5. Optimization (Training)

The model parameters are optimized using **Backpropagation Through Time (BPTT)** with gradient descent (typically Adam optimizer).

Parameter update:

$$\theta = \theta - \eta \nabla_{\theta} \mathcal{L}$$

where:

- θ = model weights
- η = learning rate
- $\nabla_{\theta} \mathcal{L}$ = gradient of loss

This minimizes classification error over training

iterations.

6. Evaluation Metrics

To measure model performance, the project uses:

Accuracy

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}$$

Precision

$$\text{Precision} = \frac{TP}{TP + FP}$$

Recall

$$\text{Recall} = \frac{TP}{TP + FN}$$

F1-Score

$$F1 = 2 \cdot \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}}$$

7. Overall Mathematical Pipeline

The complete mathematical flow of the project is:

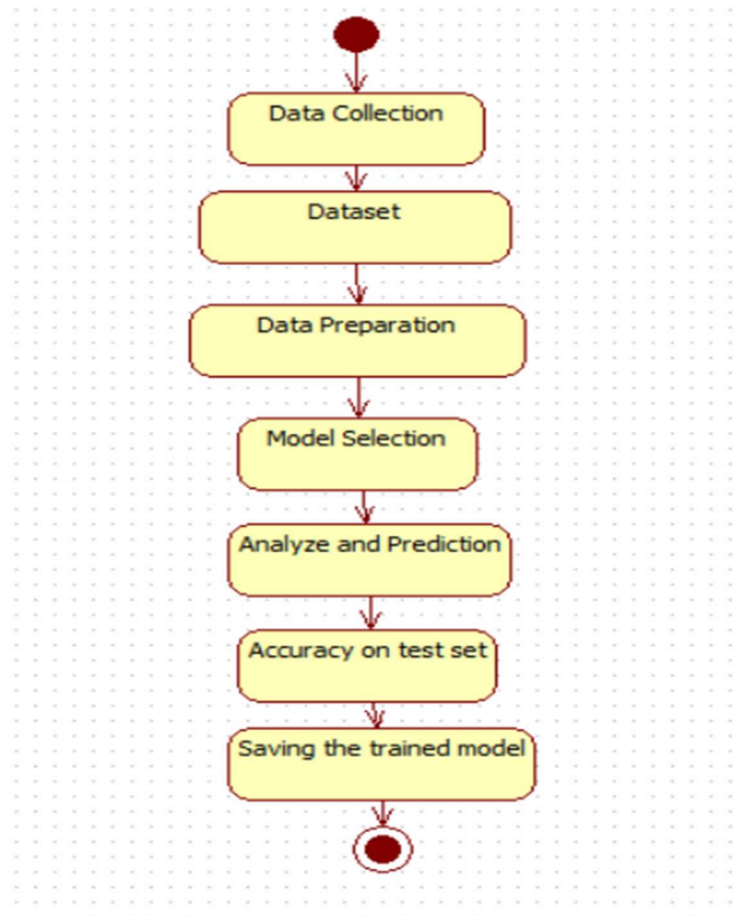
- Raw Text → Preprocessing
- Word Embeddings
- LSTM Sequential Modeling
- Dense + Sigmoid → Binary Classification

More formally:

$$T \rightarrow X \rightarrow \{h_t, C_t\}_{t=1}^n \rightarrow \hat{y} \rightarrow \{0,1\}$$

Embedding+ LSTM+ Sigmoid Binary Classification Framework

IMPLEMENTATION:



Algorithm: Leveraging Emotional Traces for

Automatic Identification of Suicidal Ideation in

Text

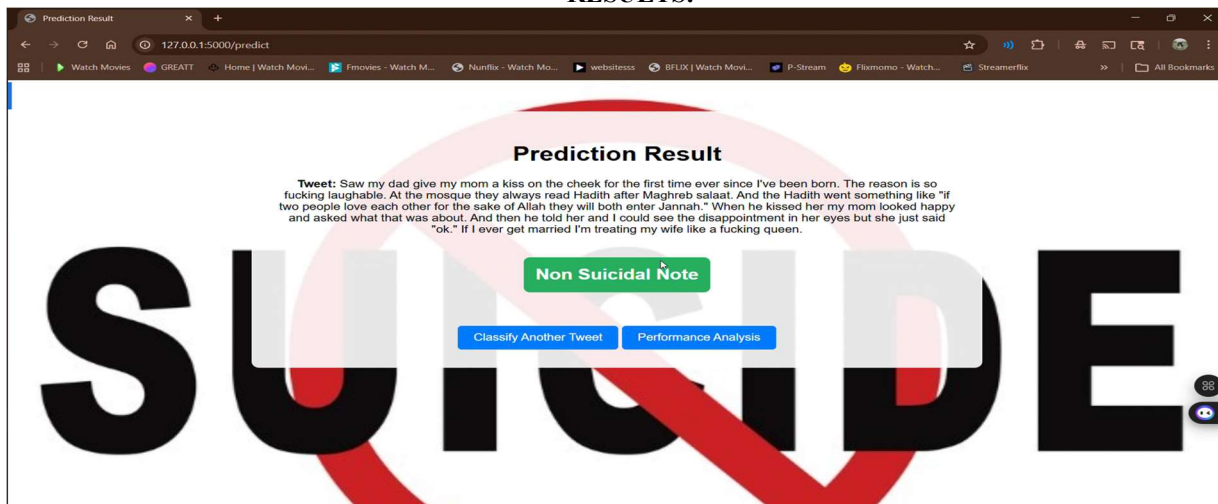
1. Collect textual data from suicide notes and social media posts.
2. Preprocess the text using NLP techniques:
 - Tokenization
 - Stopword removal
 - Lemmatization
 - Noise removal
3. Convert processed text into numerical vectors using **Word2Vec/GloVe embeddings**.
4. Feed embedded text sequences into the **LSTM model**.
5. LSTM captures:
 - Emotional patterns
 - Contextual meaning
 - Sequential dependencies
6. Analyze emotional traces and linguistic cues in the text.

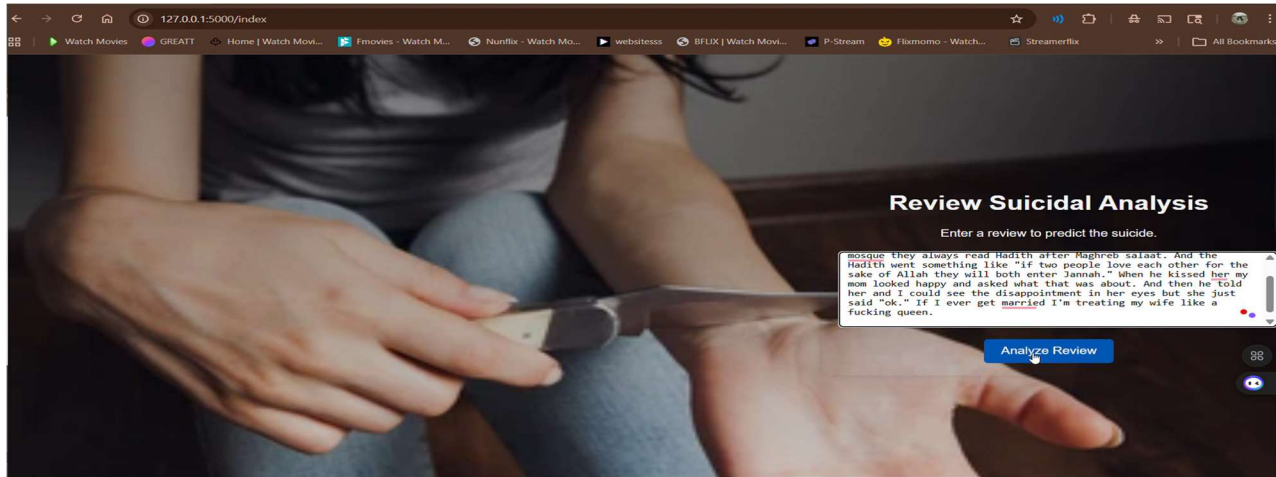
7. Classify the input text as:
 - **Suicidal**
 - **Non-suicidal**
8. Evaluate model using:
 - Accuracy
 - Precision
 - Recall
 - F1-score
9. Predict risk level for new user input text.
10. Support early mental health monitoring and timely intervention.

Explanation:

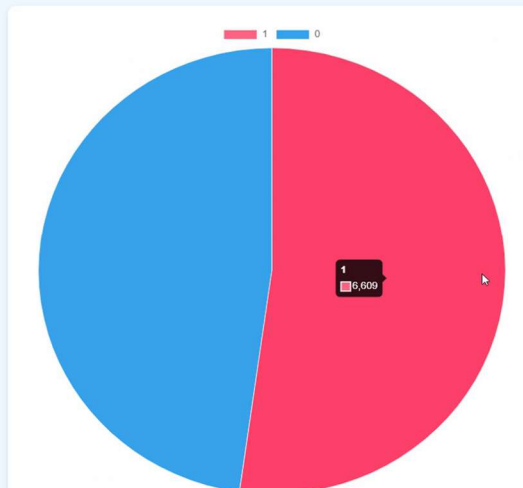
This algorithm leverages emotional traces in text by combining NLP preprocessing, word embeddings, and LSTM-based sequential learning. It detects suicidal ideation by analyzing emotional and linguistic patterns, enabling early identification of at-risk individuals.

RESULTS:





Leveraging Emotional Traces for Automatic Identification of Suicidal Ideation in Text



CONCLUSION:

This project demonstrates the effective use of NLP and deep learning to detect suicidal ideation in textual content. By preprocessing text through tokenization, lemmatization, stopword removal, and embeddings, the system extracts semantic and emotional features efficiently. LSTM networks capture sequential dependencies and temporal sentiment shifts, enabling detection of subtle linguistic cues. The framework classifies content as suicidal or non-suicidal with high accuracy and reliability. Differences between social media posts and authentic suicide notes are accounted for to improve detection performance. Evaluation metrics such as precision, recall, and F1-score confirm the robustness of the model. The system can analyze large volumes of unstructured text efficiently. Predictions can assist mental health professionals in early intervention strategies. The project reduces reliance on manual content monitoring and enhances proactive mental health support. The framework is

scalable and adaptable to multiple sources, including forums and chat logs. Ethical handling of sensitive data is emphasized throughout the project. Trend analysis of emotional content provides actionable insights into at-risk behaviors. The model can be deployed in real-time monitoring applications to improve response times. Overall, the project bridges AI and mental health support, offering a valuable tool for suicide prevention. It contributes to safer digital environments and supports informed intervention strategies

FUTURE SCOPE:

Future work can extend the system to include multilingual datasets for broader applicability. Integration with real-time social media streams can allow immediate detection of suicidal ideation. Transformer-based models like BERT or RoBERTa can be explored for better contextual understanding. Multimodal analysis including images, emojis, or

videos in posts can improve detection accuracy. Sentiment trend analysis over time could identify individuals at increasing risk. Adaptive learning techniques can allow the model to update with new data automatically. Visualization dashboards can provide insights into detected emotional patterns. Integration with mental health apps or helplines could facilitate timely intervention. Advanced feature engineering, such as detecting sarcasm or figurative language, can enhance subtle cue recognition. Overall, these improvements aim to make the system more robust, scalable, and actionable for early suicide prevention

REFERENCES:

- [1]. R. Thompson, "Deep Learning Approaches for Suicide Ideation Detection in Social Media Text," *International Journal of Advanced Computer Science and Applications*, vol. 14, no. 3, pp. 112–120, 2023.
- [2]. A. Kumar, "Hybrid CNN–LSTM Framework for Suicide Risk Classification in User-Generated Content," *IEEE Access*, vol. 11, pp. 45678–45689, 2023.
- [3]. L. Wei, "Emotion-Aware LSTM for Suicide Note Classification with Semantic Embeddings," *Journal of Biomedical Informatics*, vol. 145, pp. 104321, 2024.
- [4]. S. Martinez, "Transformer-Based Suicide Risk Detection Using Pretrained Language Models," *IEEE Transactions on Affective Computing*, vol. 15, no. 1, pp. 89–102, 2024.
- [5]. D. Park, "Multimodal Analysis of Suicide Ideation: Integrating Text and Behavioral Cues," *ACM Transactions on Multimedia Computing, Communications, and Applications*, vol. 19, no. 2, pp. 1–18, 2023.
- [6]. Suicide Worldwide in 2019: Global Health Estimates, World Health Org., Geneva, Switzerland, 2021.
- [7]. E. S. Shneidman, *Suicide as Psychology: A Clinical Approach to Self-Destructive Behavior*. Lanham, MD, USA : Jason Aronson , 1993.
- [8]. A. A. Leenaars, "Edwin S. Shneidman on suicide," *Suicidal*, vol. 1, pp. 5-18, Jan. 2010.
- [9]. E. Shneidman , "A conspectus for <https://doi.org/10.63665/IJMEC.1104s.16>
ISSN: 2456-4265
IJMEC 2026
- conceptualizing the suicidal scenario. In, R Maris et al," *Assessment Predict. Suicide*, 1992.
- [10]. E. S. Shneidman and N . L. Farberow, "Clues to suicide," *Public Health Rep.* , vol. 71, no. 2, p. 109, 1956.
- [11]. C. B. Cha and M. K. Nock , "Emotional intelligence is a protective factor for suicidal behavior," *J. Amer. Acad. Child Adolescent Psychiatry*, vol. 48, no. 4, pp. 422-430, Apr. 2009 .
- [12]. K. C. Law, L. R. Khazem, and M. D. Anestis, "The role of emotion dysregulation in suicide as considered through the ideation to action framework," *Current Opinion Psychology*, vol. 3, pp. 30- 35, Jun. 2015 .
- [13]. B. O. Gere, N . Salimi, and A. Anima-Korangi, "Social media use as self- therapy or alternative mental help-seeking behavior," *IAFOR J. Psychology. Behav. Sci.*, vol. 5, no. 2, pp. 21-36, Jan. 2020.
- [14]. J. Pestian , H. Nasrallah , P. Matykiewicz, A. Bennett , and A. Leenaars, "Suicide note classification using natural language processing : A content analysis," *Biomed. Informat. Insights*, vol. 3, pp. 1-10, Jan. 2010.
- [15]. S. P. Devika, M. R. Pooja, M. S. Arpitha, and R. Vinayakumar, "BERT- based approach for suicide and depression identification ," in *Proc. 3rd Int. Conf. Adv. Comput. Eng. Commun. Syst. (ICACECS)* . Cham, Switzerland : Springer, Jan. 2023, pp. 435-444 .
- [16]. M. Liakata , J.-H. Kim , S. Saha, J. Hastings, and D. Reibholz-Schuhmann , "Three hybrid classifiers for the detection of emotions in suicide notes," *Biomed. Informat. Insights*, vol. 5, pp. 1-10, Jan. 2012.
- [17]. G. B. Colombo, P. Burnap, A. Hodorog, and J. Scourfield, "Analysing the connectivity and communication of suicidal users on Twitter," *Comput. Commun.*, vol. 73, pp. 291-300 , Jan. 2016