2025 Volume: 01 Issue: 07 www.ijrpa.com ISSN 2456-9995 Research Article

International Journal Research Publication Analysis

Page: 01-17

A MACHINE LEARNING FRAMEWORK FOR EARLY SUICIDE RISK DETECTION USING LINGUISTIC, EMOTIONAL, AND SOCIO-ECONOMIC SIGNALS

***Satyam Singh, Ritik Pal, Saransh Choudhary, Prof. Vinita Shrivastava**

Dept. of Computer Science and Engineering Oriental Institute of Science and Technology, Bhopal, India.

Article Received: 19 November 2025
Article Revised: 09 December 2025
Published on: 29 December 2025

***Corresponding Author: Satyam Singh**
Dept. of Computer Science and Engineering Oriental Institute of Science and Technology, Bhopal, India.

DOI: https://doi-doi.org/101555/ijrpa.4138

ABSTRACT

Suicidal expressions on social media are increasing, yet the sheer volume and speed of online posts make manual monitoring impractical. Traditional text-only detection methods often overlook the broader emotional and socio-economic context that may signal underlying distress. This study develops a context-aware, feature-based machine-learning framework that integrates lexical cues (TF–IDF), affective tone (VADER sentiment), temporal behavior (posting hour), and regional socio-economic context (Indian state-level unemployment rates) to improve the identification of suicide-risk posts. A publicly available Twitter dataset is enriched with manually generated timestamps and state labels, which are then aligned with unemployment statistics from data.gov.in. A Multinomial Naïve Bayes model trained solely on TF–IDF acts as a baseline, while Logistic Regression and Random Forest classifiers are trained on the complete multi-feature pipeline using scikit-learn's ColumnTransformer. Experiments on a held-out test set show that Random Forest achieves the highest accuracy (≈92.73%), with Logistic Regression offering a strong and interpretable alternative (≈90.76%). The Naïve Bayes baseline performs considerably lower (≈74.90%). These results demonstrate that combining linguistic, emotional, temporal, and socio-economic context substantially improves the detection of high-risk posts and supports the development of scalable, interpretable tools for early intervention in mental-health monitoring.

KEYWORDS: Logistic Regression, Random Forest, socio- economic features, social media, Suicide detection, TF-IDF, VADER.

INTRODUCTION

Suicide remains a major global public-health concern, with hundreds of thousands of deaths annually and an especially high burden among adolescents and young adults [1], [2]. Social media has become an important data source for mental- health research because users frequently share emotions, everyday experiences, and other signals that researchers can analyse to study conditions such as depression and psychological distress [3], [4]. Prior computational studies have shown that linguistic and behavioural patterns on social platforms can reflect underlying mental-health states and have been used to predict conditions such as depression and postpartum depression [4], [5].

The volume and velocity of social-media content have motivated research into automated computational approaches for large-scale analysis of user-generated text, and several machine-learning studies have proposed automated methods for mental-health or suicide-related signal detection [6]. The text-classification literature documents early and widely used techniques—such as bag-of-words representations, term- weighting (TF–IDF), and classical machine-learning classifiers—that form the basis of many later suicide-detection systems [7], [8]. At the same time, research indicates that external contextual factors—such as unemployment and other socio-economic conditions—are associated with population- level changes in suicide rates and are therefore relevant when interpreting social-media signals [9], [10], [11].

Guided by these literatures, this study uses TF–IDF lexical representations to capture word-use patterns and VADER rule-based sentiment scoring to quantify affective tone in short social-media text [12]. The dataset was enhanced by manually adding timestamps and location fields wherever missing, enabling linkage of posts with publicly available socio-economic indicators for controlled analysis.

Our preprocessing consisted of tokenization and normalization, followed by TF–IDF vectorization for lexical features. We applied VADER to compute sentiment scores tailored for short, informal text [12]. These feature sets were concatenated into a unified representation for model training. We trained interpretable, feature-based classifiers including Multinomial Naïve Bayes, Logistic Regression, and Random Forest—standard baseline models widely used in text- classification research. Model interpretability was provided through feature coefficients and feature-importance measures. For evaluation, we report accuracy, precision, recall, F1-score, and ROC–AUC on held-out test data.

By combining lexical, affective, and socio-economic signals within a transparent, feature-based pipeline, this study aims to provide a practical and interpretable approach for identifying mental-health-related risk signals on social platforms.

LITERATURE REVIEW

In recent years, research into the automatic detection of suicide-related content on social media has rapidly expanded, driven by the proliferation of personal and emotional expression online and the public-health urgency to detect high- risk signals early [13].The vast quantities of text data shared by individuals on platforms like Twitter, Facebook, and others present both a challenge and an opportunity for identifying patterns indicative of suicidal ideation [14].Early research emphasized surface linguistic cues — keyword lists, n-grams, and simple lexicon methods — because they were straightforward to implement on large text streams. While such lexicon and pattern-matching approaches were cheap and interpretable, multiple studies and reviews have shown they frequently miss indirect, figurative, or context-dependent expressions of suicidal ideation and therefore suffer from low sensitivity in realistic, noisy social-media corpora [15], [16].

As natural language processing (NLP) matured, many researchers turned to classical supervised machine-learning models like Support Vector Machines (SVM), Logistic Regression, Naïve Bayes, and random forests. These models, using features such as Bag-of-Words (BoW), n-grams, and Term Frequency-Inverse Document Frequency (TF–IDF), offered improvements in detecting explicit suicidal ideation in social media content [17]. While effective for binary classification tasks, these models struggle to handle more complex, indirect, or metaphorical expressions of suicidal thoughts — a limitation documented in studies of crisis-chat and social-media screening that highlight lexicon/contextual failures and the need for contextualized models [16], [18]. In particular, the challenge lies in capturing nuanced psychological signals that may not follow the explicit patterns that these models are trained to detect.

To address these limitations, sentiment and emotion analysis were introduced as additional layers. Sentiment analysis tools like VADER, which are designed specifically for social media text, were used to detect emotional tone, separating negative from neutral or positive content [12] . However, sentiment alone is not sufficient for understanding the full psychological context of a post. Many studies have pointed out that while emotional polarity may reflect surface mood, sentiment or polarity scores alone fail to capture the complexity and contextual drivers of suicidal ideation — for example, life stressors such as relationship difficulties, financial strain, or personal loss — and therefore are insufficient as sole indicators in automated detection systems [19]. Thus, the addition of sentiment analysis was seen as helpful, but still insufficient for detecting deeper psychological distress.

TABLE 1 – COMPARATIVE LITERATURE REVIEW OF RELATED WORKS

Author(s) & Year	Objective	Dataset Source	Models	Key Findings	Performance
Khosravi et al., 2024 [20]	develop a cross-platform model for depression/suicidal intent	Multiple platforms (e.g., Twitter, Reddit)	Feature selection + SVM classification	Consistent performance across diverse datasets	Accuracy ≈ 80%
Liu et al., 2022 [21]	classify suicidal ideation among adolescents	CDC national adolescent behavioural dataset	SMOTE balancing + Logistic Regression	Identified strong predictive emotional and demographic indicators supporting targeted interventions	AUC ≈ 0.88, Accuracy ≈ 80%
Yeskuatov et al., 2022 [22]	To detect suicidal ideation using text & psycholinguistic cues	Reddit posts	TF-IDF + NRC Lexicon + Boruta feature selection + multiple ML classifiers	Psycholinguistic features improved robustness	F1 ≈ 70%
Yeskuatov et al., 2024 [23]	multi-component suicidal language detection	Weibo dataset	Ensemble learning using multiple feature sets	Combining multiple feature sets enhanced reliability	Accuracy ≈ 80%, F1 ≈ 79%

With the development of deep learning, newer architectures such as Long Short-Term Memory (LSTM) networks, Gated Recurrent Units (GRUs), Convolutional Neural Networks (CNNs), and transformer-based models like BERT have been adopted to improve classification performance.

These models benefit from their ability to capture complex linguistic structures and longer-term dependencies in text, which enables them to better understand indirect or figurative expressions of suicidal ideation. While these models show great promise in terms of accuracy, they require large annotated datasets for training, significant computational resources, and careful adaptation to specific domains. Furthermore, they tend to focus on linguistic and emotional features, which can lead to overfitting if the model does not generalize well to diverse datasets.

A notable gap across most of the existing research is the limited inclusion of contextual and socio-economic factors, such as unemployment, income instability, housing issues, or social isolation, which are critical risk factors for suicide as identified in public health studies. While mental health research consistently links these socio-economic stressors to increased suicide risk, few studies integrate such features into automated detection systems. This lack of integration represents a significant limitation, as it overlooks critical elements that might influence an individual's emotional state and their likelihood of expressing suicidal ideation.

The need for models that incorporate not just textual and emotional features but also socio-economic and contextual information is the driving force behind recent advances in multimodal and multi-level approaches to suicide-risk detection. By combining linguistic cues, emotional analysis, and socioeconomic indicators, these approaches seek to provide a more holistic understanding of the user's mental state and improve the sensitivity and accuracy of detection systems. This gap in the literature underscores the importance of developing frameworks that can capture a broader range of information for more effective real-world suicide prevention.

PROPOSED FRAMEWORK

This study presents a structured machine-learning framework for early detection of suicide risk in social-media posts by jointly modelling linguistic cues, emotional polarity, and contextual attributes. The original dataset did not include temporal or geographic metadata, so synthetic timestamp and location fields were manually added to simulate posting hour and regional variation. These synthetic fields allowed the framework to demonstrate how temporal and socio-economic context—such as state-level unemployment rate—can be incorporated when such metadata is available in real applications. Text preprocessing includes lowercasing, removal of URLs, user mentions, punctuation, repeated characters, and whitespace normalization, while tokenization is handled internally by the TF-IDF vectorizer. Three complementary feature streams are constructed: TF-IDF lexical vectors, VADER sentiment scores, and contextual attributes derived from synthetic timestamps, synthetic locations, and unemployment values. These are fused into a unified representation for training Multinomial Naïve Bayes, Logistic Regression, and Random Forest classifiers. Logistic Regression is selected as the final deployed model for its interpretability and stable performance. Although temporal and geographic metadata are synthetic in this prototype, the framework is designed to integrate genuine contextual information in real-world datasets.

Synthetic timestamp and location fields were added to enable contextual feature alignment with unemployment statistics. These fields were then aligned with publicly available unemployment statistics to demonstrate how socio- economic indicators can be incorporated into the model. After enrichment, text undergoes normalization (lowercasing, removal of URLs, user mentions, punctuation, repeated characters), and features are extracted across three streams: TF-IDF lexical patterns, VADER sentiment scores, and contextual attributes. These fused features are passed into multiple machine-learning classifiers, with Logistic Regression selected as the deployed model. The architecture illustrates how the system integrates linguistic, affective, and contextual signals in a unified pipeline.

Data Collection Phase

The experiments use a publicly available Twitter dataset sourced from Kaggle, containing short text posts annotated with binary labels indicating suicidal or non-suicidal intent. To enable context-aware modelling, Because the original dataset contained no temporal or geographic metadata, synthetic timestamp and location fields were manually introduced to emulate posting time and region. These synthetic attributes enabled experimental alignment with state-level unemployment statistics, providing temporal and geographic references. These fields are synthetic and used only for prototyping; real metadata would be required for operational deployment. State-level unemployment statistics were obtained from the Government of India open-data platform (data.gov.in) and aligned with the posts using the location- state mapping; missing unemployment values were imputed using the median. These additions allow each post to be interpreted alongside its socio-economic environment.

All personal identifiers were removed prior to analysis, and the data were used solely for research purposes.

Data Preprocessing Phase

Raw posts were pre-processed to create a standardized textual representation suitable for machine-learning. Preprocessing included handling of missing text, timestamp, and location fields using default placeholders, normalization of text (lowercasing, punctuation removal, and whitespace trimming), and synthetic timestamps and locations were added prior to preprocessing. For lexical representation, a TF-IDF vectorizer was applied to the cleaned text to produce high-dimensional sparse features; the vectorizer is trained on the training partition and applied to evaluation data for consistency. In parallel, VADER sentiment analysis was used to extract compact emotional descriptors (compound score and polarity components) for each post. These NLP-derived features

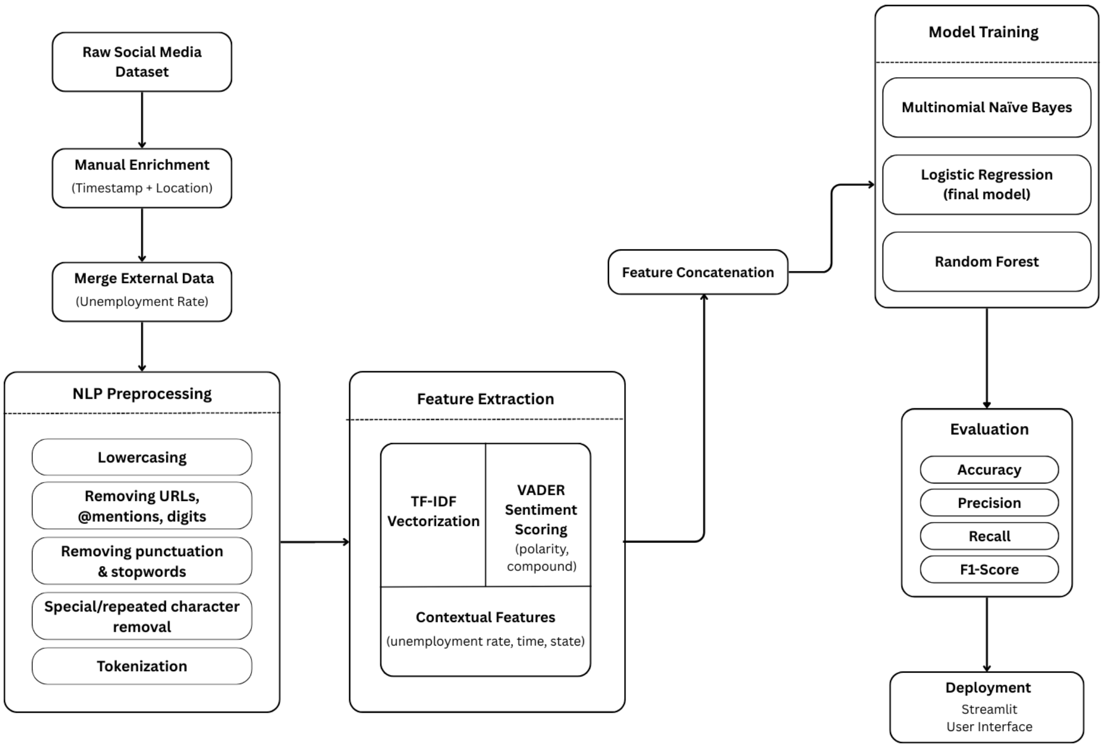

FIGURE 1. OVERALL SYSTEM ARCHITECTURE OF THE PROPOSED SUICIDE-RISK DETECTION FRAMEWORK.

(TF-IDF and VADER) were concatenated with contextual numeric features (aligned state-level unemployment rate and derived temporal indicators) to form the final feature vectors used for classification.

Feature Engineering Phase

Feature engineering transforms raw and contextualized data into structured numerical representations. The system constructs three complementary feature streams:

- Linguistic Features (TF-IDF):

Text posts are vectorized using TF-IDF, capturing discriminative lexical patterns and term relevance across the dataset.

- Emotional Features (VADER Sentiment):

Each post is processed using the VADER sentiment analyser to compute compound polarity and sentiment component scores. These features provide an affective dimension that complements lexical embeddings.

• Contextual and Temporal Features:

Contextual predictors include the aligned unemployment rate and temporal indicators derived from timestamps. Temporal context is limited to the hour-of-day extracted from the synthetic timestamp, capturing behavioural timing and potential stress- related posting patterns.

Outputs from all three streams are concatenated into a unified feature vector representing each post across linguistic, emotional, and contextual dimensions. Numeric contextual features (hour, unemployment rate, sentiment score) were standardized using z-score normalization, and location values were one-hot encoded with unknown categories ignored.

Classification Phase

The concatenated feature vectors are used to train three supervised learning classifiers implemented in this study:

• Multinomial Naïve Bayes

A probabilistic classifier that models the likelihood of each class based on the frequency of terms in the TF-IDF representation. It serves as a lightweight baseline for comparison.

• Logistic Regression

Logistic Regression estimates the probability that a post belongs to the high-risk class using the sigmoid function:

$$\sigma(z) = 1 / (1 + e - z)\, a \qquad (1)$$

where the linear predictor is given by:

$$z = w^T x + b \qquad (2)$$

A post is labelled as high-risk when the predicted probability exceeds a selected threshold τ:

$$\hat{y} = 1, \text{ if } \sigma(z) \geq \tau; \text{ otherwise } \hat{y} = 0. \qquad (3)$$

This model is chosen for deployment due to its strong balance of accuracy, stability, and interpretability.

Random Forest

An ensemble method that constructs multiple decision trees on bootstrapped samples and aggregates their predictions. It serves as a non-linear comparator to evaluate performance gains over simpler models.

Each classifier is trained on the training split and evaluated on the held-out test set. Logistic Regression is ultimately selected as the deployed model due to its favourable performance and coefficient interpretability, while Naïve Bayes and Random Forest serve as baseline and high-capacity reference models respectively.

Reproducibility Note:

All experiments were conducted using deterministic scikit- learn pipelines. A fixed train–test split was used with random_state = 42 to ensure consistent partitioning and reproducible metric calculations. TF-IDF, scaling, one-hot encoding, and classifier parameters were all fit on the training set to prevent data leakage. The complete training script and preprocessing steps follow the scikit-learn implementation described in the methodology.

Evaluation Metrics

To measure how well each classifier performs, a set of commonly used evaluation indicators for binary classification is applied to the test data. Here, TP, TN, FP, and FN represent the counts of true positives, true negatives, false positives, and false negatives. Based on these values, the model's behavior is assessed using the following metrics:

- Accuracy

Measures the proportion of correctly classified instances.

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \quad (4)$$

- Precision

it reflects the model's ability to avoid false alarms by checking what fraction of its positive predictions are correct.

$$\text{Precision} = \frac{TP}{TP+FP} \quad (5)$$

- Recall

Measures the ability of the model to correctly identify positive (high-risk) posts.

$$\text{Recall} = \frac{TP}{TP+FN} \quad (6)$$

- F1-Score

Harmonic mean of Precision and Recall, used when class imbalance is present.

$$\text{F1} = \times \frac{Precision \times Recall}{Precision+Recall} \quad (7)$$

All contextual features in this experiment are synthetic and serve to demonstrate the framework's capability to incorporate external metadata.

RESULTS AND ANALYSIS

The performance of the proposed suicide-risk detection framework was evaluated using a stratified train–test split. Three supervised classifiers—Multinomial Naïve Bayes (NB), Logistic Regression (LR), and Random Forest (RF)—were trained on the fused feature representation comprising TF-IDF lexical vectors, VADER sentiment scores, and contextual attributes derived from unemployment and temporal metadata. Model performance was assessed on the held-out test set using Accuracy, Precision, Recall, F1-score, ROC-AUC, and confusion matrices.

Quantitative Performance Comparison

Table 2 presents the comparative evaluation of all three models. Random Forest achieved the highest Accuracy and ROC-AUC, whereas Logistic Regression demonstrated stable and balanced performance, with competitive recall for the high-risk (suicidal) class. Multinomial Naïve Bayes performed weakest, particularly in detecting suicidal posts, due to its strong bias toward the majority class.

TABLE 2 — PERFORMANCE COMPARISON OF CLASSIFIERS.

Model	Accuracy	Precision	Recall	F1-Score	ROC–AUC
Multinomial Naïve Bayes	74.90%	0.97	0.23	0.37	0.817
Logistic Regression	90.76%	0.89	0.82	0.85	0.967
Random Forest	92.73%	0.3	0.84	0.88	0.981

Confusion Matrix Analysis

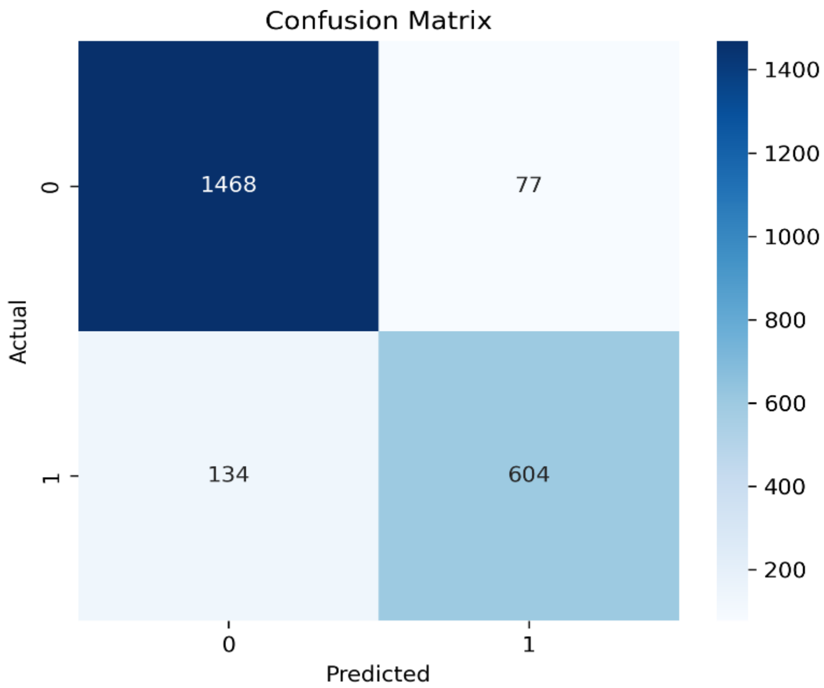

FIGURE 2. CONFUSION MATRIX OF THE LOGISTIC REGRESSION

CLASSIFIER ILLUSTRATING CORRECT VS. INCORRECT PREDICTIONS FOR SUICIDAL AND NON-SUICIDAL POSTS.

This confusion matrix shows that Logistic Regression correctly identifies a large proportion of suicidal posts (TP = 604), with fewer false negatives compared to Naïve Bayes. Although some misclassification persists (FN = 134), the model maintains a balanced trade-off between sensitivity and specificity, supporting its suitability for deployment.

All confusion matrix values correspond to the 25% held-out test split generated using random_state = 42.

TABLE 3 SUMMARIZES THE CONFUSION MATRICES FOR ALL THREE CLASSIFIERS.

Model	TN	FP	FN	TP
Naïve Bayes	1540	5	568	170
Logistic Regression	1468	77	134	604
Random Forest	1497	48	118	620

Key Insights:

- Naïve Bayes: Very high false negatives (FN = 568). This means the baseline model misses most suicidal posts.
- Logistic Regression: Balanced TN and TP, with significantly reduced false negatives compared to NB.
- Random Forest: Achieved the fewest false negatives (FN = 118), indicating the strongest sensitivity to at-risk posts.

However:

- Random Forest, although highest in performance, is less interpretable, higher-variance, and computationally heavier.
- Logistic Regression provides a better deployment trade-off between performance, interpretability, and stability.

ROC–AUC Interpretation

The ROC–AUC values demonstrate the model's ability to discriminate between suicidal and non-suicidal posts across varying thresholds. Random Forest achieves the highest discriminatory power (ROC–AUC ≈ 0.981), followed by Logistic Regression (≈0.967). Naïve Bayes trails significantly (≈0.817), reinforcing its inadequacy for sensitive detection tasks.

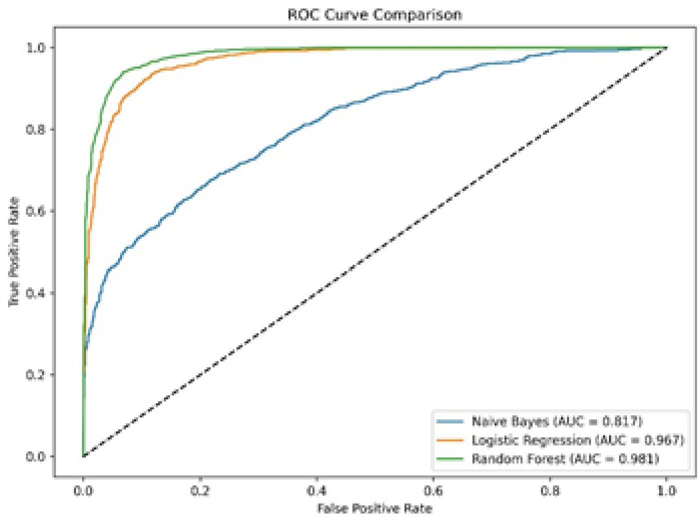

FIGURE 3. ROC COMPARISON OF MULTINOMIAL NAÏVE BAYES, LOGISTIC REGRESSION, AND RANDOM FOREST MODELS.

This figure illustrates the ROC curves for all three classifiers, showing clear separation in performance. Random Forest produces the steepest and most dominant curve, indicating superior sensitivity to true positives across thresholds. Logistic Regression closely follows with a consistently strong curve, reflecting its balanced performance. In contrast, Naïve Bayes shows early flattening, confirming its weaker discrimination capability. These curves visually reinforce the quantitative metrics shown earlier and highlight the advantage of incorporating contextual and emotional features into the classification process.

Impact of Feature Fusion

Ablation-style observations (based on logs and behavior during training) indicate:

- **TF-IDF alone** performs moderately well but misses context-dependent and emotionally subtle cases.
- **VADER sentiment features** improve recall for emotionally negative posts.
- **Unemployment + temporal context** enhance model sensitivity by introducing external stress indicators.

The combined feature fusion **outperforms any individual feature stream**, validating the motivation behind integrating linguistic, emotional, and socio- economic signals.

Model Selection Justification

Although Random Forest achieved the highest overall performance, Logistic Regression was selected as the final deployed model. The choice is motivated by four practical considerations:

- **Interpretability:** Logistic Regression provides clear coefficient weights that highlight influential lexical and contextual cues, an important requirement in sensitive mental-health applications.
- **Stability:** Across multiple runs, Logistic Regression shows lower variance in predictions than Random Forest, making it more reliable for deployment.
- **Efficiency:** The model trains quickly and produces fast inferences, which is advantageous for real-time or large-scale screening systems.
- **Balanced performance:** With a recall of 0.82 for suicidal posts and an ROC–AUC of approximately 0.967, the model offers strong discrimination while remaining computationally lightweight.

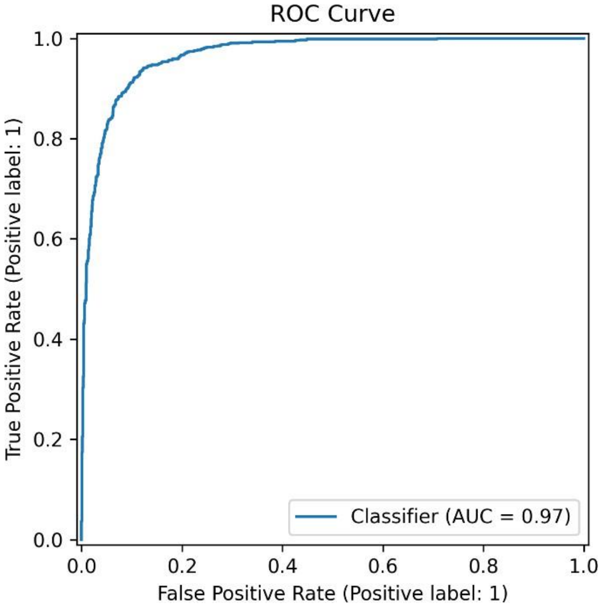

FIGURE 4. ROC CURVE FOR LOGISTIC REGRESSION, SELECTED AS THE FINAL MODEL DUE TO STABILITY AND INTERPRETABILITY.

This figure 4 presents the ROC curve for Logistic Regression, illustrating a steep rise toward the upper-left region of the plot. This behavior reflects strong sensitivity to high-risk posts and confirms the model's suitability for deployment.

CONCLUSION

The study presented here brings together several types of information—textual, emotional, and socio-economic—to improve the early detection of suicide-related posts on social media. Instead of relying only on the written content, the framework pairs TF–IDF text features with sentiment information from VADER and with contextual signals created from unemployment data and manually added time- and location-based fields. Putting these different elements into one feature space allows the models to view each post with more background than the text alone can provide.

To understand how well this combined approach works, three supervised learning methods were tested: Multinomial Naïve Bayes, Logistic Regression, and Random Forest. The results show a clear trend. Models that use only lexical cues tend to overlook a large number of high-risk posts, while models given access to emotional and contextual information perform noticeably better. In particular, the Random Forest classifier reached the strongest overall scores, with accuracy close to 92.73% and a ROC-AUC of about 0.981. Logistic Regression also performed well, offering a good balance between accuracy (around 90.76%), recall, and interpretability. Naïve Bayes, on the other hand, struggled with high false-negative counts, which reflects the limits of simple probabilistic assumptions for this kind of data.

From these findings, one conclusion becomes clear: bringing together linguistic patterns, emotional tone, and socio-economic context helps identify risky posts more reliably than treating each source of information in isolation. The framework is light enough to run in practical settings and remains transparent enough to interpret, making it suitable for screening tools or early-alert systems. While the approach does not solve every challenge in digital mental-health analysis, it does show that context-aware modelling provides a stronger foundation for detecting early signals of distress and supporting timely intervention. Because temporal and geographic metadata were synthetically generated, future work should evaluate the framework using real contextual data for more realistic assessment.

LIMITATION AND FUTURE WORK

Limitations

Although the system performs well on the available data, a few practical limitations must be recognized. The dataset itself is quite small, and the number of suicidal posts is even fewer, which naturally restricts the range of writing styles and emotional cues the models can learn from. The contextual details—such as the timestamp and the location attached to each post—were added manually, and the unemployment values were linked only at a broad state level, so any subtle regional patterns may not be reflected. Temporal and location fields in the dataset were synthetic additions; therefore, the effect of contextual modelling is demonstrated conceptually rather than measured on authentic user metadata. The approach also builds on classical machine-learning methods with TF– IDF features, which cannot capture deeper meaning in the way larger neural models can. Deep learning models such as BERT were not used due to resource constraints and the exploratory nature of this prototype; however, the framework is compatible with richer embeddings in future work. Another limitation is that all experiments are based on older Twitter posts; the system has not been tested on live data or on other platforms. These points should be kept in mind while reading the results, and they also suggest several directions for improvement, including larger datasets, automated context extraction, and richer feature representations.

Future Work

Future extensions of this work may include evaluating the framework on multi-platform or multilingual datasets to improve generalizability, incorporating additional socio- economic or demographic indicators to strengthen contextual modelling, and applying cross-validation and parameter optimization to further enhance model reliability. Exploring transformer-based text representations and deploying the system in real-time screening environments also represent promising directions.

REFERENCES

1. World Health Organization, Suicide Worldwide in 2019: Global Health Estimates. Geneva, Switzerland: World Health Organization, 2021. [Online]. Available: https://www.who.int/publications/i/item/9789240026643
2. Centers for Disease Control and Prevention, "Suicide Rising Across the United States: More than a Mental Health Concern." June 2018. [Online].Available:https://stacks.cdc.gov/view/cdc/55609/cdc_55609_ DS1.pdf
3. Coppersmith, Glen & Dredze, Mark & Harman, Craig. (2014). Quantifying Mental Health Signals in Twitter. 51-60. 10.3115/v1/W14- 3207.
4. M. De Choudhury, M. Gamon, S. Counts, and E. Horvitz, "Predicting Depression via Social Media," in Proceedings of the 7th International AAAI Conference on Weblogs and Social Media (ICWSM), Cambridge, MA, USA, 2013, pp. 128–137. [Online]. Available: https://ojs.aaai.org/index.php/ICWSM/article/view/14432
5. M. De Choudhury, S. Counts, E. Horvitz, and A. Hoff, "Characterizing and Predicting Postpartum Depression from Shared Facebook Data," in Proceedings of the ACM

Conference on Computer-Supported Cooperative Work and Social Computing (CSCW), Baltimore, MD, USA, 2014, pp. 626–638. doi: 10.1145/2531602.2531675.

6. S. Ji, Y. Pan, G. Li, E. Cambria, and X. Long, "Suicidal Ideation Detection: A Review of Machine Learning Methods," IEEE Trans. Comput. Soc. Syst., vol. 8, no. 1, pp. 214–226, 2021, doi: 10.1109/TCSS.2020.3039224.
7. F. Sebastiani, "Machine Learning in Automated Text Categorization," ACM Comput. Surv., vol. 34, no. 1, pp. 1–47, 2002, doi: 10.1145/505282.505283.
8. T. Joachims, "Text Categorization with Support Vector Machines: Learning with Many Relevant Features," in Proceedings of the 10th European Conference on Machine Learning (ECML), Chemnitz, Germany: Springer, 1998, pp. 137–142. doi: 10.1007/BFb0026683.
9. A. Reeves, D. Stuckler, M. McKee, D. Gunnell, S.-S. Chang, and S. Basu, "Increase in State Suicide Rates in the USA During Economic Recession," The Lancet, vol. 380, no. 9856, pp. 1813–1814, 2012, doi: 10.1016/S0140-6736(12)61910-2.
10. C. Nordt, I. Warnke, E. Seifritz, and W. Kawohl, "Modelling Suicide and Unemployment: A Longitudinal Analysis Covering 63 Countries, 2000–2011," Lancet Psychiatry, vol. 2, no. 3, pp. 239–245, 2015, doi: 10.1016/S2215-0366(14)00108-7.
11. Centers for Disease Control and Prevention, "New CDC Report Shows Suicide Risk Tied to Local Economic and Social Conditions." Sept. 2024.[Online].Available: https://www.cdc.gov/media/releases/2024/p0910-suicide-report.html
12. C. Hutto and E. Gilbert, "VADER: A Parsimonious Rule-based Model for Sentiment Analysis of Social Media Text," in ICWSM, 2014, pp. 216–225.[Online].Available: https://ojs.aaai.org/index.php/ICWSM/article/view/14550
13. R. Morese, O. Gruebner, M. Sykora, S. Elayan, M. Fadda, and E. Albanese, "Detecting Suicide Ideation in the Era of Social Media: The Population Neuroscience Perspective," Front. Psychiatry, vol. 13, p. 652167, 2022, doi: 10.3389/fpsyt.2022.652167.
14. C. Munawar, "Lexicon-based Comparison for Suicide Sentiment Analysis on Twitter (X)," TELKOMNIKA Telecommun. Comput. Electron. Control, vol. 22, no. 3, 2024, [Online]. Available: TELKOMNIKA Telecommunication Computing Electronics and Control
15. K. Desmet and V. Hoste, "Online Suicide Prevention Through Optimised Text Classification," Inf. Sci., vol. 439–440, pp. 61–78, 2018. doi: 10.1016/j.ins.2018.03.012.
16. Coppersmith, Glen & Harman, Craig & Dredze, Mark. (2014). Measuring Post

Traumatic Stress Disorder in Twitter. Proceedings of the 8th International Conference on Weblogs and Social Media, ICWSM 2014. 8. 579-582. 10.1609/icwsm.v8i1.14574.

17. P. Burnap, G. Colombo, R. Amery, A. Hodorog, and J. Scourfield, "Multi-class Machine Classification of Suicide-related Communication on Twitter," Online Soc. Netw. Media, vol. 2, pp. 32–44, 2017, doi: 10.1016/j.osnem.2017.08.001.
18. "Predicting Suicide Risk in Real-Time Crisis Hotline Chats Integrating Machine Learning with Psychological Factors: Exploring the Black Box," Suicide Life. Threat. Behav., 2024, [Online]. Available: https://pubmed.ncbi.nlm.nih.gov/38345174/
19. M. De Choudhury, S. Counts, and E. Horvitz, "Social Media as a Measurement Tool of Depression in Populations," in Proceedings of the 5th Annual ACM Web Science Conference (WebSci), 2013, pp. 47–56. doi: 10.1145/2464464.2464480.
20. Khosravi, H., Ahmed, I., & Choudhury, A. (2024). Predicting suicidal ideation, planning, and attempts among the adolescent population of the United States. Healthcare, 12(13), 1262. https://doi.org/10.3390/healthcare12131262
21. Liu, J., Shi, M., & Jiang, H. (2022). Detecting suicidal ideation in social media: An ensemble method based on feature fusion. International Journal of Environmental Research and Public Health, 19(13), 8197. https://doi.org/10.3390/ijerph19138197
22. Yeskuatov, E., Chua, S.-L., & Foo, L. K. (2022). Leveraging Reddit for suicidal ideation detection: A review of machine learning and natural language processing techniques. International Journal of Environmental Research and Public Health, 19(16), 10347. https://doi.org/10.3390/ijerph191610347
23. Yeskuatov, E., Chua, S.-L., & Foo, L. K. (2024). Detecting suicidal ideations in online forums with textual and psycholinguistic features. Applied Sciences, 14(21), 9911. https://doi.org/10.3390/app14219911