



Detection of Suicidal Intent in Tweets based on Machine Learning and Natural Language Processing

Md Asadul Islam¹, Nusrat Jahan^{2*}, Sumiya Haque³, Md Shawon Ali⁴

^{1, 3, 4} Department of Computer Science and Engineering, Sonargaon University, Dhaka, Bangladesh

^{2*} Department of International Business, Ajou University, Suwon, Republic of Korea

*E-mail: sumiyaque107@gmail.com

Received: April 27, 2026

Revised: May 30, 2026

Accepted: May 29, 2026

Keywords:

Suicide Detection, Twitter, Machine Learning, NLP, TF-IDF, Mental Health

ABSTRACT

In this research, detection of suicidal intent in tweets based on machine learning and natural language processing (NLP) is considered to be an early detection system for vulnerable individuals and to facilitate early intervention. Social media websites are also a popular medium for the expression of feelings like distress and suicide, and therefore detection with the support of automation is the starting point towards monitoring mental health. The article employs an open-source collection of 9,119 pre-processed tweets that are labelled as suicidal or non-suicidal, removing special characters, emails, and tokenization to give the data cleanliness. Text data were numerically represented using TF-IDF vectorization in which unigram and bigram features were preserved in order to facilitate correct training of the model. Four supervised models, including Logistic Regression, Naive Bayes, Decision Tree, and Linear Support Vector Classifier (Linear SVC), were trained on 80% of data and evaluated on 20% of data using accuracy, precision, recall, and F1-score measures. Confusion matrices were calculated to determine true positives, false positives, and other outcomes of the classifications across all models, and a straightforward demo prediction was added to test tweets to illustrate real-world application. Among models tested, Linear SVC performed best overall at 92%, followed by the Logistic Regression model, then the Naive Bayes and the Decision Tree model with lower though still superior performance. Ethical concerns were noted, including anonymization, privacy, and handling sensitive data. The study presents a reproducible and feasible method for auto-detection of suicidal tweets from Twitter and proves the effectiveness of NLP and machine learning methods in supporting mental health interventions, reducing harm, and facilitating timely intervention of vulnerable groups.

1.0 INTRODUCTION

Social networking platforms have become an essential part of modern communication, where people actively share their thoughts, emotions, and daily life experiences. Among these platforms, Twitter is widely used for posting short, emotionally expressive messages that often reflect personal struggles, distress, and in some cases, suicidal ideation (Kitchen et al., 2025). While social media provides opportunities for communication and emotional support, it can also reveal hidden psychological distress signals. If these signals are not identified in time, they may lead to severe outcomes such as self-harm or suicide. Therefore, early detection of suicidal intent from social media content is an important and timely research area with significant real-world impact. Suicide is a major global public health problem and accounts for a large number of preventable deaths every year. Traditional suicide prevention approaches such as clinical interviews, counseling, and helplines depend heavily on self-reporting and direct human interaction, which may not always be available or accessible for at-risk individuals (Kim et al., 2025). In contrast, social media platforms generate massive amounts of user-created data that can be analyzed using computational methods. This opens new opportunities to complement traditional approaches with automated detection systems.

Recent advancements in Machine Learning (ML) and Natural Language Processing (NLP) have shown strong potential in analyzing textual data for mental health applications. NLP techniques such as tokenization, TF-IDF vectorization, and n-gram modeling convert unstructured text into structured numerical representations suitable for machine learning models. These features allow algorithms to identify patterns associated with suicidal ideation. Commonly used classifiers such as Logistic Regression, Naive Bayes, Decision Trees, and Support Vector Machines have been successfully applied in sentiment analysis, cyberbullying detection, and mental health prediction tasks. These computational approaches provide scalable and cost-effective solutions for identifying at-risk individuals from social media data (Belouali et al., 2025). Despite these advantages, several challenges still exist. Social media language is highly informal and includes slang, abbreviations, emojis, and irregular sentence structures, which make text processing difficult. Tweets are often very short and may lack sufficient context, making accurate interpretation challenging.

In addition, ethical issues such as user privacy, informed consent, and responsible handling of sensitive data are extremely important in this domain. Without proper safeguards, automated systems may lead to misclassification or misuse of personal information. Therefore, strong data preprocessing, anonymization, and secure handling mechanisms are necessary.

The main motivation of this study is to develop a reliable and reproducible framework for detecting suicidal intent from tweets using machine learning and NLP techniques. This research contributes to existing work by evaluating multiple classification models, comparing their performance, and demonstrating their applicability in real-world scenarios (Hussen et al., 2026). Using a publicly available dataset of suicidal and non-suicidal tweets, the study investigates how different models perform in terms of accuracy, precision, recall, and F1-score, and how preprocessing and feature extraction affect classification results. The problem addressed in this research is the automatic detection of suicidal ideation from short social media text. Specifically, the study focuses on identifying which machine learning model performs best for this task, how preprocessing and feature engineering impact model performance, and what ethical considerations must be taken into account when deploying such systems in real-world environments.

The objectives of this study are as follows: (1) to preprocess and clean Twitter data by removing noise such as special characters, URLs, and irrelevant content while preserving meaningful information; (2) to extract textual features using TF-IDF and n-gram techniques for numerical representation; (3) to train and evaluate multiple supervised machine learning models including Logistic Regression, Naive Bayes, Decision Tree, and Linear Support Vector Classifier (LinearSVC); (4) to analyze model performance using evaluation metrics such as accuracy, precision, recall, F1-score, and confusion matrix; and (5) to address ethical concerns including privacy protection, data anonymization, and responsible use of automated detection systems (Mamun et al., 2025). The scope of this study is limited to English-language tweets with binary classification labels (suicidal and non-suicidal). It does not include other social media platforms, multimodal data such as images or videos, or long-form text analysis.

Additionally, the model is intended for supportive and preventive purposes rather than clinical diagnosis. Any real-world application must involve collaboration with mental health professionals to ensure ethical and safe usage. In summary, this introduction highlights the importance of suicide detection from social media, outlines the research motivation, problem statement, objectives, and limitations, and establishes a foundation for the subsequent sections of the study, including methodology, experiments, and results analysis in the field of computational mental health.

2.0 LITERATURE REVIEW

Suicidality detection from social media is an emerging research area due to the high prevalence of mental health issues and the extensive use of platforms such as Twitter, Facebook, and Instagram. Literature shows that individuals often express distress, hopelessness, and suicidal thoughts through tweets and posts, providing opportunities for computational techniques to identify vulnerable users. Automated systems attempt to utilize these signals for early intervention, offering scalable support alongside traditional mental health screening (Hsin et al., 2025). Natural Language Processing (NLP) plays a key role in processing social media text by converting unstructured data into structured features for machine learning models. Common preprocessing steps include tokenization, removal of special characters, stopword removal, and text normalization through lowercasing or stemming. TF-IDF vectorization is widely used to measure word importance in a corpus, while n-gram models capture sequential word context. These techniques transform text into a format suitable for identifying suicidal ideation patterns. Machine learning models are widely used for suicidal and non-suicidal classification (Abdelmoteleb, Ghallab, & IsHak, 2025).

Logistic Regression provides an efficient binary classification method with interpretable coefficients reflecting feature importance (Sreevalsan-Nair et al., 2025). Naive Bayes is simple and effective for text classification but assumes feature independence. Decision Trees enable hierarchical decision-making and capture non-linear relationships, though they may overfit if not properly controlled (Lestandy et al., 2025). Linear Support Vector Classifier (LinearSVC) performs well in high-dimensional text space by separating classes using hyperplanes.

Model performance varies depending on preprocessing, feature representation, and dataset characteristics, making careful model selection necessary. Evaluation metrics are essential for assessing model performance. Accuracy measures overall correctness, precision evaluates correctly identified suicidal posts, recall measures detection of actual suicidal posts, and F1-score balances precision and recall. Confusion matrices provide detailed insight into true positives, false positives, true negatives, and false negatives for error analysis.

Model explainability is also important in mental health applications to ensure predictions are interpretable and reliable.

Social media language is highly informal, containing slang, abbreviations, emoticons, and irregular grammar, which makes traditional NLP processing difficult. Tweets are often short and lack context, and their meaning may depend on external or prior information. In addition, labeled datasets are often limited or biased, which may affect model performance. Ethical issues such as privacy, consent, and responsible data handling are critical in this domain and require careful consideration. Further challenges include cross-platform generalization and the need for multimodal analysis involving text, images, or videos. Feature extraction techniques such as sentiment analysis, part-of-speech tagging, and keyword extraction help detect implicit emotional cues. Ensemble methods combining multiple classifiers have shown strong performance in improving precision-recall balance. Real-time monitoring systems have also been developed for continuous analysis of social media streams for early intervention. Integration of automated systems with professional mental health support is increasingly recognized as a useful approach, enabling early detection while maintaining ethical standards and user confidentiality.

Overall, existing literature confirms that suicide detection through social media is a complex task enabled by NLP and machine learning techniques. Models such as Logistic Regression, Naive Bayes, Decision Trees, and LinearSVC are commonly used, with performance evaluated using precision, recall, F1-score, and confusion matrices. However, challenges such as informal language, contextual dependency, imbalanced datasets, and ethical concerns still remain. Addressing these issues requires the development of standardized, interpretable, and ethical frameworks for large-scale mental health intervention. This review serves as the foundation

for the next sections of the research, including dataset preparation, methodology, and model evaluation for suicidal tweet detection.

3.0 DATASET DESCRIPTION AND ETHICAL ISSUES

The dataset used in this research consists of publicly available English tweets collected from Twitter for suicidal ideation detection. Twitter is widely used for expressing emotions and personal experiences, making it a suitable platform for analyzing mental health-related text data. The dataset contains a total of 9,119 tweets, labeled into two categories: suicidal (1) and non-suicidal (0). It includes 3,998 suicidal tweets and 5,121 non-suicidal tweets, ensuring a balanced distribution for supervised learning. Data collection followed ethical guidelines for publicly available social media research. Tweets were gathered using keyword-based searching related to suicidal behavior and mental health expressions. Only public posts were included, while private or protected content was excluded to respect user privacy and consent. Personally identifiable information such as user IDs was removed to ensure anonymity. Preprocessing was applied to clean and normalize the text data. Unnecessary elements such as URLs, email addresses, punctuation, and special characters were removed to reduce noise. All text was converted to lowercase, and tokenization was performed to split sentences into individual words. Stopwords were retained selectively to preserve emotional and contextual meaning relevant to suicidal expression. TF-IDF vectorization was then applied to convert text into numerical feature representations. Additionally, unigram and bigram features were used to capture contextual word patterns in short social media text.

Ethical considerations were strictly followed throughout the dataset preparation. All data was anonymized, and no attempt was made to identify users. The dataset was used only for research purposes, and sensitive content was handled carefully. It is important to note that the dataset is intended for computational analysis only and not for clinical diagnosis. Proper evaluation metrics such as precision, recall, F1-score, and confusion matrices were used to ensure reliable and responsible model assessment. Figure 1: illustrates the class distribution of the dataset, showing the number of suicidal (1) versus non-suicidal (0) tweets used for model training and evaluation

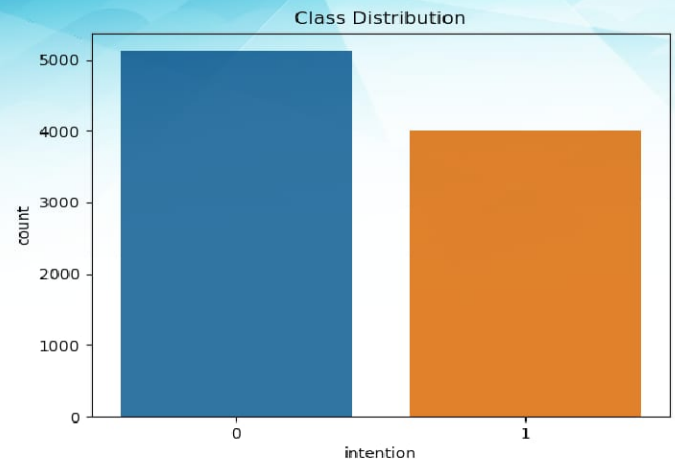


Figure 1: Class Distribution Plot

4.0 METHODOLOGY

The proposed methodology consists of a sequence of interdependent steps, including data preprocessing, feature extraction, train-test splitting, model training, and evaluation. These steps are designed to ensure efficiency, accuracy, and reproducibility in detecting suicidal ideation from Twitter data (Weridin & Wyss, 2024).

4.1 Data Pre-processing

Data preprocessing is a fundamental stage in handling raw and noisy social media text. Twitter data is highly unstructured and often contains irrelevant or redundant elements that can negatively affect model performance. Therefore, the dataset was carefully cleaned by removing URLs, email addresses, user mentions, punctuation marks, and special characters. These elements do not contribute to semantic meaning and may introduce noise during feature extraction. All text was converted into lowercase to ensure uniformity and reduce redundancy caused by case sensitivity. Tokenization was then applied to split each sentence into individual words, making the text suitable for numerical transformation. Because social media language is highly informal, additional preprocessing considerations were required. This includes handling slang terms, abbreviations, and repeated characters commonly found in tweets. Stopword handling was performed selectively. While traditional NLP pipelines often remove stopwords, this study retained certain words when they contributed to emotional or contextual meaning. For example, words that help express intensity or emotional state were preserved when relevant to suicidal ideation detection. The final output of this stage was a clean, normalized, and structured dataset ready for feature extraction.

4.2 Feature Extraction

After preprocessing, the textual data was converted into numerical form using TF-IDF (Term Frequency– Inverse Document Frequency). TF-IDF is a widely used technique in text classification that assigns higher weights to words that are important within a document while reducing the influence of frequently occurring common words across the dataset (Liao et al., 2024). This ensures that meaningful terms related to suicidal ideation receive higher importance in the model. To enhance contextual understanding, both unigram and bigram features were used. Unigrams capture individual words, while bigrams capture pairs of consecutive words. This is important because suicidal expressions often appear in phrases rather than isolated words. For example, phrases such as “feel hopeless,” “want die,” or “no reason live” carry stronger meaning than single words alone. The final feature space was limited to 5,000 dimensions to maintain a balance between performance and computational efficiency. This dimensionality reduction helps reduce overfitting while preserving the most informative linguistic patterns (Ivey-Stephenson, 2024). The resulting feature matrix serves as input for machine learning models.

4.3 Model Selection and Training

Four supervised machine learning models were selected for classification: Logistic Regression, Naive Bayes, Decision Tree, and Linear Support Vector Classifier (LinearSVC). These models were chosen to represent different learning paradigms, including probabilistic, linear, and tree-based approaches. Logistic Regression is a linear model used for binary classification. It is simple, efficient, and provides interpretable results by estimating the probability of class membership. It performs well with sparse text data generated from TF-IDF features. Naive Bayes is a probabilistic classifier based on Bayes’ theorem. It assumes feature independence and is widely used in text classification due to its speed and effectiveness. Despite its simplicity, it often performs well in high-dimensional datasets. Decision Tree is a non-linear model that splits data into subsets based on feature conditions. It can capture complex relationships between features but is prone to overfitting. To prevent this, constraints such as maximum depth were applied during training. LinearSVC is a Support Vector Machine-based classifier that performs well in high-dimensional spaces. It finds an optimal hyperplane that separates classes with maximum margin. It is particularly

effective for sparse text data and is known for strong performance in classification tasks involving TF-IDF representations. All models were trained using the processed dataset with appropriate hyperparameter tuning to improve performance. The training process focused on maximizing predictive accuracy while maintaining generalization ability and avoiding overfitting.

4.4 Train-Test Split

The dataset was divided into training and testing sets using an 80:20 ratios. The training set was used to build and optimize the models, while the testing set was used to evaluate their performance on unseen data. This ensures unbiased evaluation and measures how well the model generalizes to new inputs. Stratified sampling was applied during splitting to maintain equal representation of suicidal and non-suicidal classes in both sets. This is important because class imbalance can lead to biased models that favor the majority class. Stratification ensures fair evaluation across both categories.

4.5 Evaluation Metrics

Model performance was assessed using standard evaluation metrics: accuracy, precision, recall, and F1-score. Accuracy measures the overall proportion of correct predictions. Precision evaluates the proportion of correctly identified suicidal tweets among all tweets predicted as suicidal. Recall measures the ability of the model to correctly identify all actual suicidal tweets. F1-score provides a balanced measure between precision and recall. In addition, confusion matrices were used to analyze model predictions in detail. They provide a breakdown of true positives, false positives, true negatives, and false negatives. This helps in understanding classification errors and identifying areas for improvement in model performance. Overall, this methodology integrates preprocessing, feature extraction, multiple machine learning models, structured data splitting, and comprehensive evaluation techniques to build a reliable framework for suicidal ideation detection from social media text.

5.0 RESULTS AND DISCUSSION

This section evaluates the effectiveness of supervised machine learning techniques in classifying suicidal tweet content. Four models Logistic Regression, Naive Bayes, Decision Tree, and

Linear Support Vector Classifier (LinearSVC) were trained on an 80% data split and validated on the remaining 20%. The performance was quantified using accuracy, precision,

recall, and F1-score, providing a clear summary of model strengths for real-world mental health surveillance.

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
LinearSVC	92.16	91.68	89.39	90.52
Logistic Regression	90.68	93.29	83.77	88.27
Naive Bayes	88.93	85.03	89.27	87.09
Decision Tree	88.05	89.35	80.76	84.83

5.1 Model Performance Comparison

As shown in Table 1, performance varied significantly across the different classification paradigms. While accuracy provides a general overview, the precision-recall trade-off is more critical in the context of sensitive suicidal content detection.

Table 1: Comparative Performance Metrics for Suicidal Intent Detection

5.2 Analytical Insights

vbitions:

- **Linear SVC:** Achieved the highest overall performance. Its high recall (89.39%) and precision (91.68%) indicate that it is equally adept at identifying genuine suicidal posts while minimizing false alarms. This stability makes it the most suitable model for high-dimensional TF-IDF feature spaces.
- **Logistic Regression:** Demonstrated the highest precision (93.29%). However, its lower recall (83.77%) suggests a "cautious" classification approach, which reduces false positives but increases the risk of missing vulnerable individuals.
- **Naive Bayes:** While maintaining a strong recall (89.27%), it suffered from poor precision (85.03%). This indicates a tendency to over-predict suicidal intent, which could lead to high false-alarm rates in monitoring systems.

- **Decision Tree:** Although inherently interpretable, this model showed the poorest recall (80.76%). Its struggle with sparse, high-dimensional data often leads to under-estimating high-risk cases, making it less reliable for this specific dataset.

The confusion matrices further confirmed these patterns, highlighting how LinearSVC successfully balances the sensitivity-specificity trade-off compared to the over-estimation seen in Naive Bayes or the under-estimation in Decision Trees.

Figure 2: illustrates a bar chart comparing model performance metrics such as Accuracy, Precision, Recall, and F1-score for all four classifiers.

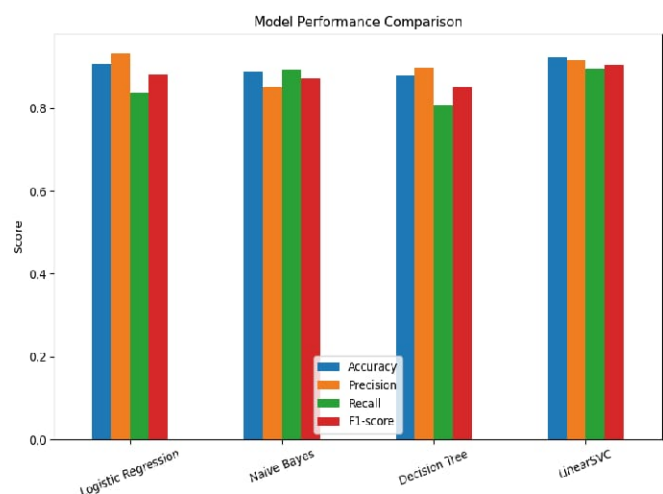


Figure 2: Model Performance Comparison Bar Chart

5.3 Predictive Analysis

Additional testing of test tweet model performance once more reveals true performance. For instance, the sentence "I feel so hopeless and want to give up" was rightly classified as suicidal (1) by Logistic Regression, Naive Bayes, and LinearSVC but wrongly classified as non-suicidal (0) by Decision Tree. On the other hand, a case of a non-suicidal positive sample such as "I am having a good day, everything is wonderful!" was correctly identified as non-suicidal (0) by all except Decision Tree, which performed steadily but dropped in sensitivity. That deviation indicates that linear models (LinearSVC and Logistic Regression) are better under the sparsity of high-dimensional text data created through TF-IDF vectorization. Probabilistic nature of Naive Bayes ensures it to obtain very high recall but could over-predict and sensitivity and specificity will have to be sacrificed. Decision Trees, although explainable and non-linear pattern detection enabled, could have a little lower performance on sparse and high-dimensional data prevalent in short social media updates.

Figure 3: illustrates sample tweet predictions, displaying example tweets with their predicted labels (0 for non-suicidal, 1 for suicidal) generated by the trained models.

```

--- Logistic Regression ---
I feel so hopeless and want to give up. -> 1
I am having a great day, everything is awesome! -> 0

--- Naive Bayes ---
I feel so hopeless and want to give up. -> 1
I am having a great day, everything is awesome! -> 0

```

Figure 3: Demo Predictions Screenshot

5.4 Error Analysis

Error analysis provides important insights into the limitations and improvement areas of the models. False negatives (suicidal posts classified as non-suicidal) are the most critical errors because they may delay necessary intervention. LinearSVC and Logistic Regression produced fewer false negatives compared to other models, although no model completely eliminated them. False positives were more common in Naive Bayes, which may lead to unnecessary alerts or concern. Misclassifications mainly occurred in tweets with vague expressions, sarcasm, or implicit emotional distress. Very short tweets with limited context were also

difficult to classify correctly. In some cases, cultural expressions and idiomatic language further reduced accuracy. These results suggest that richer contextual understanding, such as embeddings or transformer-based models, could improve performance in future work.

6.0 KEY INSIGHTS

The results highlight that LinearSVC provides the best balance between precision and recall, making it suitable for monitoring systems. Naive Bayes performs well in recall but increases false positives, while Decision Tree offers interpretability but lower stability. This shows a clear trade-off between performance and interpretability. TF-IDF representation and proper text cleaning significantly improve classification performance. Stratified train-test splitting also ensures fair evaluation by maintaining class balance. Finally, ethical responsibility is essential when interpreting results. These models should support mental health analysis, not replace clinical diagnosis. Misclassifications must be handled carefully, and results should always be used with human oversight and professional support.

6.1 Practical Implications

The findings suggest that LinearSVC and similar models can be integrated into early warning systems for detecting suicidal content in social media. However, combining automated systems with human review is necessary to ensure both accuracy and ethical safety. Continuous improvement using error feedback can further enhance system performance.

6.2 Case Study

A simple case study was conducted using sample tweets to evaluate model behavior. Two tweets were selected: "I feel so hopeless and want to give up" (suicidal) and "I am having a great day, everything is awesome!" (non-suicidal). After preprocessing and TF-IDF transformation, all models were tested. Logistic Regression, Naive Bayes, and LinearSVC correctly classified both tweets. However, Decision Tree misclassified the suicidal tweet, showing limitations in handling sparse text data. This comparison shows that linear and probabilistic models are more reliable for such tasks, while Decision Trees may struggle with contextual meaning.

6.3 Ethical Issues

Ethical considerations are essential in this research domain. Only publicly available data was used, and all personal identifiers were removed during preprocessing to ensure anonymity and privacy. The model outputs are not clinical diagnoses and should not be treated as final judgments. Misclassifications may have serious consequences, so results must be interpreted carefully. These systems are designed to support mental health professionals, not replace them. Due to the informal and ambiguous nature of social media language, models may still misinterpret slang, sarcasm, or emotional expressions. Therefore, human supervision, ethical deployment, and continuous monitoring are necessary to ensure safe and responsible use of automated suicide detection systems.

7.0 CONCLUSION

This study presents an end-to-end framework for suicidal tweet detection using text preprocessing, TF-IDF feature extraction, and supervised machine learning models. Four classifiers Logistic Regression, Naive Bayes, Decision Tree, and LinearSVC were evaluated using accuracy, precision, recall, F1-score, and confusion matrix analysis. Among the models, LinearSVC achieved the best performance due to its effectiveness in handling high-dimensional sparse text data and maintaining a strong balance between precision and recall. Logistic Regression also showed stable and interpretable results, making it suitable for binary classification tasks. Naive Bayes performed well in identifying suicidal cases with higher recall but produced more false positives. Decision Tree provided interpretability but was less stable compared to linear models. Overall, the study demonstrates a structured and effective approach for classifying suicidal content in social media text.

REFERENCES

A. Belouali, C. Kitchen, A. Zirikly, P. Nestadt, H. C. Wilcox, and H. Kharrazi, "Identifying and characterizing suicide decedent subtypes using deep embedded clustering," *Scientific Reports*, vol. 15, no. 1, p. 23069, 2025.

A. Z. Ivey-Stephenson, "CDC guidance for community response to suicide clusters, United States, 2024," *MMWR Supplements*, vol. 73, 2024.

C. H. Liao, C. S. Chang, P. T. Kung, W. Y. Chou, and W. Y. Tsai, "Stroke and suicide among people with severe mental illnesses," *Scientific Reports*, vol. 14, no. 1, p. 4991, 2024.

C. Kitchen, A. Zirikly, A. Belouali, H. Kharrazi, P. Nestadt, and H. C. Wilcox, "Suicide death prediction using the Maryland suicide data warehouse: A sensitivity analysis," *Archives of Suicide Research*, vol. 29, no. 2, pp. 453-467, 2025.

H. Hsin et al., "Predicting and Preventing Suicide at Entry to Mental Health Care: A Community Engaged, Machine Learning Model Implementation," *medRxiv*, Mar. 2025.

J. Hughes, B. Foley, C. Colohan, and D. Lyness, "Understanding Suicide, Drug and Alcohol Deaths in Northern Ireland: Socio-economic and Household Insights (2011-2022)," *International Journal of Population Data Science*, vol. 10, no. 4, 2025.

J. Hussen, M. Y. Mia, M. Kohily, S. Halim, and I. Hossen, "Bangla Suicidal Ideation Detection: Performance and Efficiency Benchmark of Simple and Complex Classifiers on Social Media Data," *Journal of Applied Technology and Innovation (JATI)*, vol. 10, no. 1, Apr. 2026. [Online]. Available: <https://doi.org/10.65136/jati.v10i1.11>

J. Lim et al., "The relative toxicity of medicines detected after poisoning suicide deaths in Australia, 2013–19: a data linkage case series study," *Medical Journal of Australia*, vol. 222, no. 7, pp. 339-347, 2025.

J. Sreevalsan-Nair et al., "Mental healthcare in the times of climate change action and data science," in *Data-Driven Insights and Analytics for Measurable Sustainable Development Goals*, Morgan Kaufmann, 2025, pp. 59-81.

L. C. Chen, Y. M. Bai, S. J. Tsai, C. M. Cheng, and M. H. Chen, "Eating disorders, psychiatric comorbidities, and suicide," *Journal of Affective Disorders*, 2025.

M. A. Mamun, F. Al-Mamun, M. E. Hasan, N. Roy, M. M. ALmerab, D. Gozal, and M. S. Hossain, "Exploring suicidal thoughts among prospective university students: a study with applications of machine learning and GIS techniques," *BMC Psychiatry*, vol. 25, no. 1, p. 755, 2025.

M. Lestandy, A. Abdurrahim, A. Faruq, and M. Irfan, "A comparative analysis of transfer learning models on suicide and non-suicide textual data," *TELKOMNIKA*, vol. 23, no. 2, pp. 426-434, 2025.

N. Idaikkadar et al., "Advancing Ethical Considerations for Data Science in Injury and Violence Prevention," *Public Health Reports*, vol. 140, 2025. [Online]. Available: [doi:10.1177/00333549241312055](https://doi.org/10.1177/00333549241312055).

S. Abdelmoteleb, M. Ghallab, and W. W. IsHak, "Evaluating the ability of artificial intelligence to predict suicide: A systematic review of reviews," *Journal of Affective Disorders*, 2025.

S. Kim, K. H. Jeong, D. Song, H. J. Cho, and Y. Kim, "The influence of search volume for suicide on suicide rates: focusing on gender differences," *Journal of Men's Health*, vol. 21, no. 6, pp. 108-116, 2025.

S. Werdin and K. Wyss, "Challenges in the evaluation of suicide prevention measures and quality of suicide data in Germany, Austria, and Switzerland: findings from qualitative expert interviews," *BMC Public Health*, vol. 24, no. 1, p. 2209, 2024.

