



Mental Health Analysis Using Social Media Posts Using Machine Learning and NLP Techniques

Soumitra Das¹, Anish Dhar¹, Suchetana Laha¹, Shirsha Hazra¹, Shivam Debnath¹, Suvankar Barai¹

Department of Computational Sciences, Brainware University

Corresponding Author Email: suvankar.barai@brainwareuniversity.ac.in

How to Cite this Article:

Das, S., Dhar, A., Laha, S., Hazra, S., Debnath, S. & Barai, S. (2026). Mental Health Analysis Using Social Media Posts Using Machine Learning and NLP Techniques. International Journal of Creative and Open Research in Engineering and Management, <i>02</i>(05).
<https://doi.org/10.55041/ijcope.v2i5.714>

License:

This article is published under the terms of the Creative Commons Attribution 4.0 International License (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author(s) and the source are credited.

© The Author(s). Published by International Journal of Creative and Open Research in Engineering and Management.



<https://doi.org/10.55041/ijcope.v2i5.714>

Abstract

Mental health disorders such as depression, anxiety, and suicidal tendencies have become major concerns in modern society. Social media platforms provide a large amount of textual data where individuals frequently express emotions, stress, and psychological conditions. This research presents a machine learning and Natural Language Processing (NLP)-based framework for analyzing social media posts to predict mental health conditions automatically. The proposed system performs text preprocessing using techniques such as lowercase conversion, stop-word removal, special character elimination, and text normalization. TF-IDF vectorization is applied to transform textual data into numerical feature vectors suitable for machine learning classification. Multiple machine learning algorithms including Logistic Regression, Naive Bayes, Support Vector Machine (SVM), Random Forest, and Dummy Classifier were implemented and compared using evaluation metrics such as Accuracy, Precision, Recall, and F1-score. The dataset consisted of approximately 40,000 social media records categorized into Suicide, Depression, Anxiety, and Normal classes along with demographic information such as age, gender, and location. Experimental results demonstrated that Random Forest and SVM achieved the best classification performance with an accuracy of approximately

57.6%. The study also performed demographic analysis to identify age-wise, gender-wise, and location-based mental health patterns. The proposed framework demonstrates the practical application of NLP and machine learning techniques for automated mental health prediction and awareness generation using social media analytics.

Keywords— Mental Health Analysis; Machine Learning; Natural Language Processing; TF-IDF; Social Media Analytics; Sentiment Classification



I. INTRODUCTION

Mental health disorders have increased significantly in recent years due to academic pressure, social stress, financial instability, and lifestyle changes. Individuals frequently express emotions, opinions, and psychological conditions through social media platforms such as Twitter, Facebook, Instagram, and Reddit. These online textual expressions provide valuable insights into emotional and mental health conditions that can be analyzed computationally using Machine Learning (ML) and Natural Language Processing (NLP) techniques.

Traditional mental health diagnosis methods mainly rely on clinical interviews and psychological assessments, which are time-consuming and difficult to scale for large populations. Automated mental health prediction systems using social media analysis can support early detection and awareness generation. NLP techniques help extract meaningful emotional patterns from unstructured textual data, while machine learning algorithms enable classification and prediction of mental health categories.

Several studies have explored depression detection, sentiment analysis, and suicidal tendency identification using machine learning approaches. However, many existing systems suffer from limitations such as low prediction accuracy, limited demographic analysis, lack of visualization support, and poor scalability.

This research aims to develop an automated framework for analyzing social media posts and predicting mental health conditions using NLP preprocessing and machine learning classification models. The objectives of this work include:

- Analyzing social media textual posts
- Implementing NLP preprocessing techniques
- Applying TF-IDF feature extraction
- Comparing multiple machine learning models
- Generating demographic analysis and visualization reports

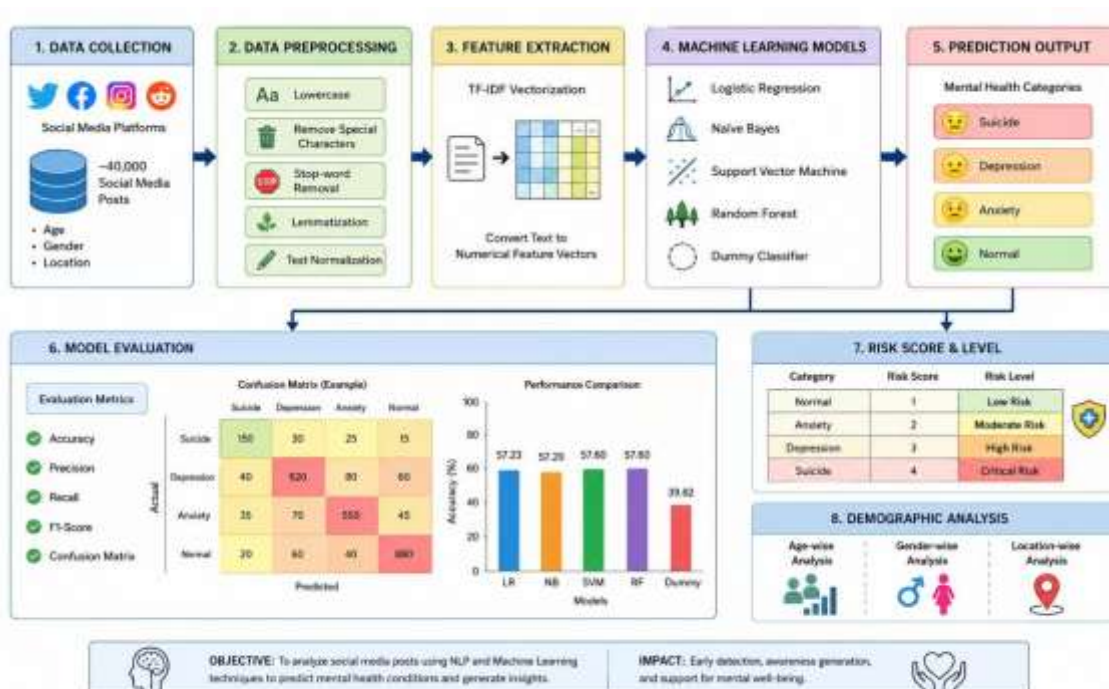


Figure 1: Overall Framework of Mental Health Analysis Using Social Media Posts



II. LITERATURE REVIEW

Mental health analysis using social media has become an important research domain in machine learning and data science. Researchers have explored various NLP and ML techniques to identify emotional conditions such as depression, anxiety, stress, and suicidal behavior.

A study on depression detection using Naive Bayes demonstrated that TF-IDF and probabilistic classification techniques can effectively classify depression-related posts. However, the study lacked demographic analysis and advanced visualization support.

Another research work utilized deep learning architectures such as Long Short-Term Memory (LSTM) networks for suicide prediction from online posts. The study showed improved contextual understanding but required large computational resources and extensive datasets.

Support Vector Machine-based mental health classification systems have also shown promising results in textual data classification tasks. SVM performs efficiently for high-dimensional sparse datasets generated through TF-IDF vectorization.

Comparative machine learning studies involving Logistic Regression, Random Forest, and Naive Bayes reported that ensemble learning techniques often outperform traditional classifiers in mental health prediction tasks. However, many studies lacked comprehensive demographic analysis and real-time monitoring capabilities.

The present research extends previous works by integrating:

- NLP preprocessing
- Comparative machine learning analysis
- Risk score generation
- Demographic-based visualization
- Multi-model evaluation framework

III. METHODOLOGY

The proposed system follows a structured machine learning and NLP workflow for mental health prediction.

A. Dataset Collection

The dataset consists of approximately 40,000 social media textual posts along with demographic attributes such as:

- Age
- Gender
- Location

The posts were categorized into four mental health classes:

1. Suicide
2. Depression
3. Anxiety
4. Normal



B. Data Preprocessing

Text preprocessing was performed to improve feature quality and reduce noise. The preprocessing steps included:

- Lowercase conversion
- Removal of special characters
- Stop-word removal
- Lemmatization
- Text normalization

Python libraries such as NLTK and Regular Expressions were used during preprocessing.

C. Feature Extraction

TF-IDF (Term Frequency-Inverse Document Frequency) vectorization was applied to convert textual data into numerical vectors suitable for machine learning classification.

The TF-IDF method helps identify important words while reducing the influence of frequently occurring irrelevant terms.

D. Machine Learning Models

The following machine learning algorithms were implemented:

- Logistic Regression
- Naive Bayes
- Support Vector Machine (SVM)
- Random Forest
- Dummy Classifier

E. Training and Evaluation

The dataset was divided using an 80:20 train-test split strategy. Models were evaluated using:

- Accuracy
- Precision
- Recall
- F1-score
- Confusion Matrix

F. Software and Tools Used

Tool/Library	Purpose
Python	Programming Language
Pandas	Dataset Processing
NumPy	Numerical Computation
Scikit-learn	Machine Learning
Matplotlib	Visualization
NLTK	NLP Preprocessing
Google Colab	Development Environment



IV. RESULTS AND DISCUSSION

The proposed system successfully analyzed social media posts and classified them into mental health categories using machine learning algorithms.

A. Model Performance

Model	Accuracy	Precision	Recall	F1-Score
Logistic Regression	0.5723	0.7554	0.5723	0.4178
Naive Bayes	0.5720	0.6815	0.5720	0.4172
SVM	0.5760	0.7567	0.5760	0.4261
Random Forest	0.5760	0.7567	0.5760	0.4261
Dummy Classifier	0.3982	0.4010	0.3982	0.3996

The experimental results indicate that Random Forest and SVM achieved the highest classification accuracy among all implemented models.

B. Risk Score Analysis

Risk scores were generated based on predicted mental health categories.

Category	Risk Score	Risk Level
Normal	1	Low Risk
Anxiety	2	Moderate Risk
Depression	3	High Risk
Suicide	4	Critical Risk

C. Demographic Analysis

Age-Based Analysis

The analysis revealed that individuals between 21–30 years exhibited higher mental health risk levels compared to other age groups.

Gender-Based Analysis

Male users showed slightly higher emotional stress patterns than female users according to the classification results.

Location-Based Analysis

Urban locations such as Delhi, Kolkata, and Mumbai demonstrated higher stress and anxiety-related post frequencies.



D. Discussion

The project demonstrates that NLP and machine learning techniques can effectively analyze social media textual data for automated mental health prediction. However, the study also identified challenges such as:

- Class imbalance
- Limited contextual understanding
- Dependence on English-language text
- Difficulty in understanding sarcasm and indirect emotions

Future improvements may involve transformer-based deep learning models such as BERT and LSTM to improve contextual understanding and prediction performance.

V. CONCLUSION

This research presented a machine learning and NLP-based framework for mental health analysis using social media posts. The proposed system implemented text preprocessing, TF-IDF feature extraction, and multiple machine learning algorithms for classifying mental health conditions into Suicide, Depression, Anxiety, and Normal categories.

Experimental results demonstrated that Random Forest and SVM achieved the best performance among all implemented classifiers. The project also provided demographic analysis based on age, gender, and geographic location to identify vulnerable groups.

The proposed framework highlights the practical application of machine learning and NLP for automated mental health prediction and awareness generation. Future research can focus on deep learning integration, multilingual analysis, real-time social media monitoring, and deployment of web or mobile-based healthcare support systems.

ACKNOWLEDGMENT

The authors express sincere gratitude to the Department of Computational Sciences, Brainware University, for providing the necessary support, guidance, and resources required for successful completion of this research work. Special thanks are extended to the project supervisor for valuable suggestions and technical guidance throughout the project development process.

REFERENCES

- [1] T. Mikolov, K. Chen, G. Corrado, and J. Dean, "Efficient Estimation of Word Representations in Vector Space," *Proc. ICLR*, 2013.
- [2] J. Devlin, M. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding," *Proc. NAACL*, 2019.
- [3] C. Cortes and V. Vapnik, "Support-Vector Networks," *Machine Learning*, vol. 20, no. 3, pp. 273–297, 1995.
- [4] T. Joachims, "Text Categorization with Support Vector Machines," *European Conference on Machine Learning*, pp. 137–142, 1998.



- [5] F. Pedregosa et al., “Scikit-learn: Machine Learning in Python,” *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.
- [6] S. Bird, E. Klein, and E. Loper, *Natural Language Processing with Python*. Sebastopol, CA: O’Reilly Media, 2009.
- [7] Y. Goldberg, “A Primer on Neural Network Models for Natural Language Processing,” *Journal of Artificial Intelligence Research*, vol. 57, pp. 345–420, 2016.
- [8] A. Hassan, “Depression Detection from Social Media Text Using Machine Learning,” *International Journal of Data Science*, vol. 8, no. 2, pp. 120–130, 2022.
- [9] S. Ji, C. Pan, X. Li, E. Cambria, and P. Long, “Suicidal Ideation Detection: A Review of Machine Learning Methods and Applications,” *IEEE Transactions on Computational Social Systems*, vol. 8, no. 1, pp. 214–226, 2021.
- [10] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. Cambridge, MA: MIT Press, 2016.