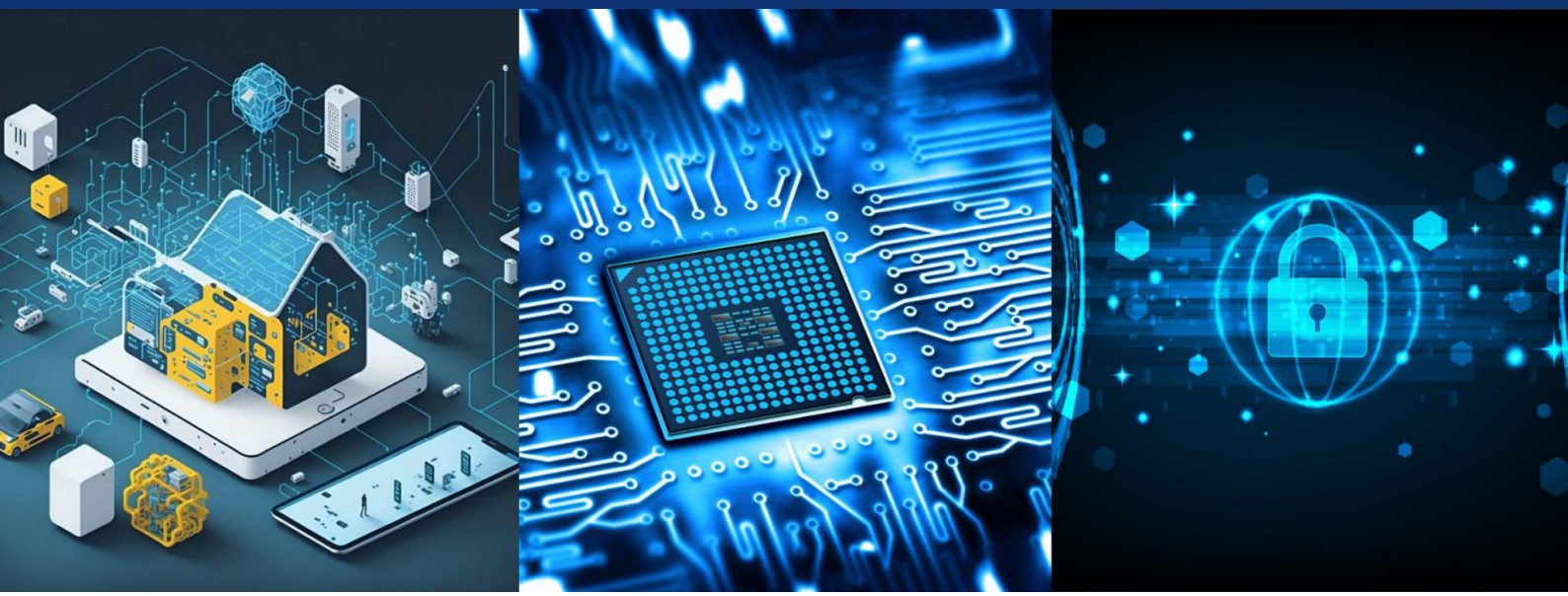
IJIRCCE

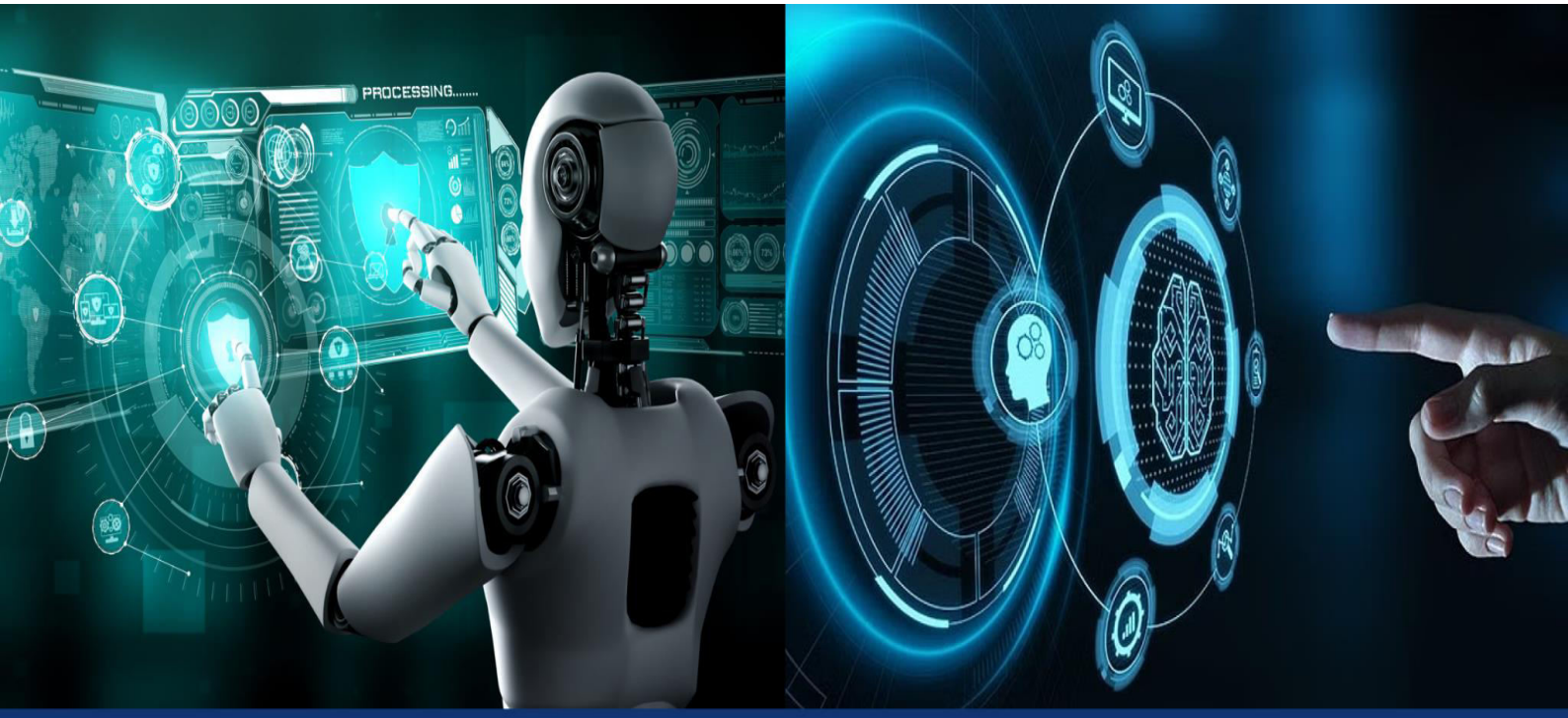
ISSN(O): 2320-9801

ISSN(P): 2320-9798



International Journal of Innovative Research in Computer and Communication Engineering

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



Impact Factor: 8.771

Volume 13, Issue 5, May 2025

© 2025 IJIRCCE | Volume 13, Issue 5, May 2025| DOI:10.15680/IJIRCCE.2025.1305140

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.771| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

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

Depressive and Non - Depressive Tweets using Sequential Deep Learning Model

Indhumathi. S, Abishek. J, Aswin .M, Dhivin .C, Gokulavasan. R

Assistant Professor, Dept. of CSE, KGISL Institute of Technology, Coimbatore, Tamil Nadu, India

UG Student, Dept. of CSE, KGISL Institute of Technology, Coimbatore, Tamil Nadu, India

ABSTRACT: With the growing influence of social media on personal expression, detecting early signs of mental health conditions through online behavior has emerged as a crucial area of research. Recognizing depressive tendencies in user-generated content, particularly tweets, offers a timely and scalable approach to support mental health awareness and intervention. This project proposes an AI-driven framework for identifying depressive tendencies in social media posts, specifically tweets, by leveraging state-of-the-art Natural Language Processing (NLP) and deep learning techniques. The system integrates advanced text preprocessing, contextual embedding models such as BERTweet and RoBERTa, sequential models like BiLSTM, and baseline classifiers including Logistic Regression to accurately distinguish between depressive and non-depressive content. Further, the solution employs a user-friendly interface built using Flask and NGROK, facilitating real-time analysis and feedback delivery. The inclusion of live data collection through Twitter API and performance evaluation via standard classification metrics ensures the model's robustness and practical relevance.

KEYWORDS: Depression Prediction, Social Media-Based Mental Health Insights, Neural Networks for Psychological Analysis, Tweet-Based Text Classification, BERTweet Model, BiLSTM Architecture, RoBERTa Language Model, Emotion Detection using NLP, AI-Driven Mental Health Evaluation, Live Mood Detection, Logistic Regression Classifier, Twitter Sentiment Analysis, Deep Learning Frameworks, Real-Time Tweet Processing, Flask Web Interface.

I. INTRODUCTION

As global focus on mental health continues to grow, using technology to identify emotional struggles at an early stage has become both relevant and necessary. Social media platforms, particularly Twitter, have become public diaries where users openly share their thoughts and emotions. Even in short form, social media content can reveal significant emotional patterns that help in identifying potential symptoms of depression. Conventional methods of assessing mental health typically involve manual evaluation and struggle to keep pace with the rapid flow and sheer volume of online content.

This project introduces an AI-powered framework designed to automatically classify tweets as depressive or non-depressive using cutting-edge deep learning and natural language processing (NLP) techniques. The system integrates intelligent text preprocessing, context-aware language models (such as BERTweet and RoBERTa), sequential modeling through BiLSTM, and baseline comparison using Logistic Regression. To ensure usability and accessibility, a lightweight Flask-based web interface supported by NGROK has been developed, enabling real-time interaction and feedback.

II.RELATED WORKS

Recent advancements have significantly contributed to tackling the challenge of detecting depressive content on social media platforms. Research by Orabi et al. provides an extensive evaluation of deep learning methods applied to sentiment analysis and depression identification. Their study highlights that Long Short-Term Memory (LSTM) networks are particularly effective in capturing sequential patterns and emotional subtleties within short texts.

Additionally, developments in transformer-based approaches, notably Nguyen et al.'s work on BERTweet, emphasize the effectiveness of using pre-trained models customized for the unique characteristics of social media text. Likewise,

© 2025 IJIRCCE | Volume 13, Issue 5, May 2025| DOI:10.15680/IJIRCCE.2025.1305140

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.771| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

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

Ji et al.'s research utilizing BiLSTM networks demonstrates the advantages of bidirectional learning in capturing deeper contextual meaning and detecting linguistic markers related to depressive behavior.

Moreover, Singh et al.'s investigation into hybrid models that integrate traditional machine learning with deep learning techniques highlights the effectiveness of ensemble methods in improving the robustness and reliability of emotional state classification from tweets. Together, these varied approaches provide a deeper insight into the technological advancements for detecting depressive content, forming a strong basis for developing more accurate, scalable, and real-time mental health monitoring solutions through social media platforms..

III.PROPOSED ALGORITHM

A. Design Considerations:

- Dataset consists of tweets labeled as depressive and non-depressive.
- Preprocessing includes stopword removal, stemming, and text normalization.
- Features are extracted using TF-IDF vectorization.
- Multiple classifiers are used: SVM, Logistic Regression, Random Forest, and XGBoost.
- Ensemble voting strategy is applied to improve accuracy.
- Accuracy, Precision, Recall, and F1-score are used as evaluation metrics.
- Real-time classification support is considered for future integration.

B. Description of the Proposed Algorithm:
The goal of the proposed system is to classify tweets as depressive or non-depressive using deep learning and ensemble machine learning techniques to increase the accuracy and robustness of mental health analysis.

Step 1: Data Preprocessing
Each tweet undergoes a sequence of preprocessing tasks:
Tokenization and lowercasing.
Removal of URLs, mentions (@user), hashtags, and special characters.
Stopword removal and stemming.
Transformation into numerical feature vectors using TF-IDF.

Step 2: Model Training and Prediction
Four machine learning models are trained:
SVM (Support Vector Machine): Best suited for text classification with high dimensional data.
Logistic Regression: For binary classification baseline.
Random Forest: For handling noise and improving generalization.
XGBoost: Gradient-boosted decision trees to optimize classification accuracy.
Each model is trained on the same preprocessed dataset. The individual predictions are then collected for ensemble voting.

Step 3: Ensemble Classification using Soft Voting
The final prediction is made using Soft Voting:
Each model returns the probability of the tweet being depressive.
Probabilities are averaged.
The class with the highest average probability is selected.
This method ensures better generalization and reduces the likelihood of bias from any single model.

IV.RESULTS AND DISCUSSION

This discussion highlights the strengths, challenges, and areas for improvement in the approach used to classify depressive and non-depressive tweets through deep learning techniques.

Effectiveness: The combination of modern Natural Language Processing (NLP) methods with sequential deep learning models like LSTM, BiLSTM, and BERT variants has greatly enhanced the accuracy of detecting mental health

indicators in tweets. Additionally, the system's ability to process tweets in real-time or batches enables ongoing analysis of mental health trends, supporting early detection and response by mental health practitioners.

Challenges: The informal nature of tweets—including slang, abbreviations, and multilingual content—adds complexity to text preprocessing and model training. Moreover, the reliance on annotated datasets limits scalability, and ethical concerns such as user privacy and data sensitivity must be carefully managed.

Opportunities: As deep learning and NLP technologies continue to evolve, there is considerable potential to improve the system's reliability and adaptability. Enhancements such as incorporating emotion-focused embeddings, supporting multiple languages, and utilizing large-scale unsupervised learning can boost detection accuracy across varied populations. When implemented responsibly with strong data privacy measures, these models offer promising applications in real-world mental health monitoring and broader public health initiatives.

Future Directions: The development of lightweight models for mobile deployment and integration with mental health chatbot interfaces can make the technology more accessible. Additionally, collaboration between AI researchers, psychologists, and policy makers is crucial to ensure responsible usage and real-world impact.

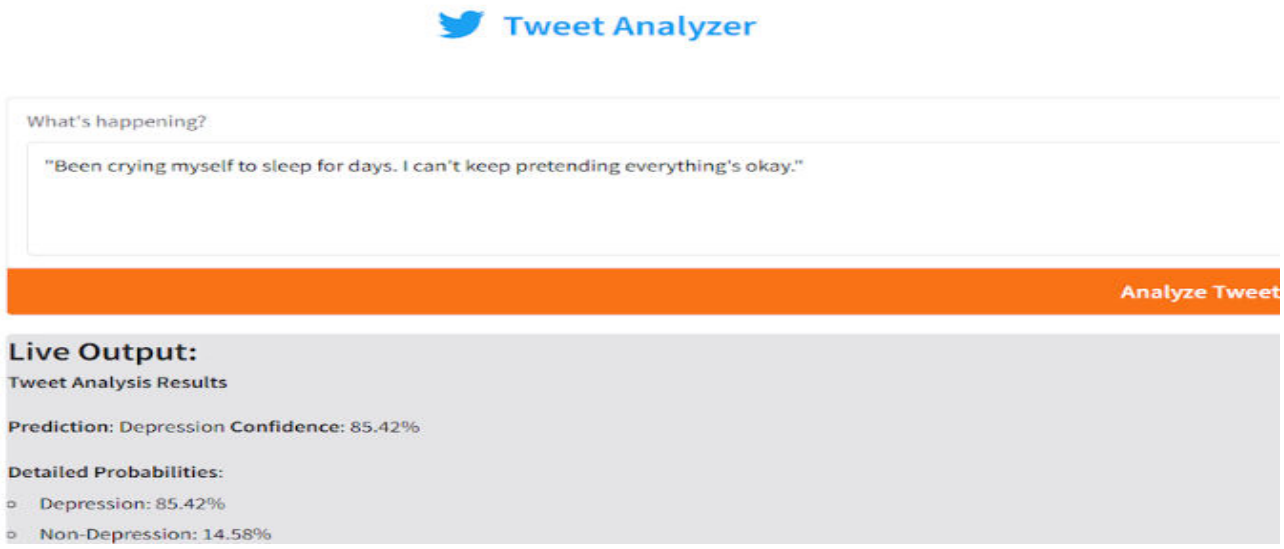


Figure.1

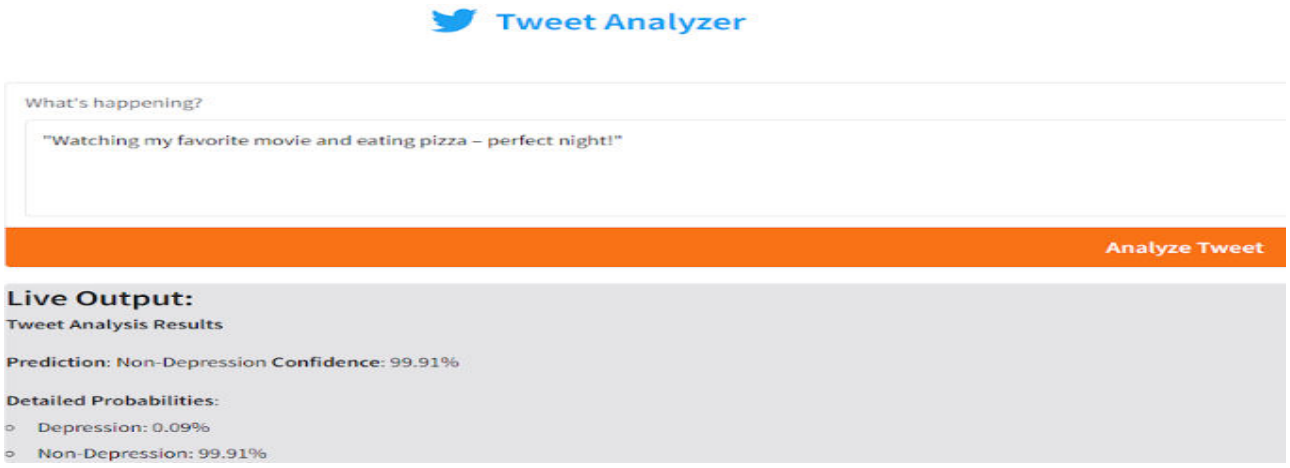


Figure.2

© 2025 IJIRCCE | Volume 13, Issue 5, May 2025| DOI:10.15680/IJIRCCE.2025.1305140

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.771| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

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

V. CONCLUSION AND FUTURE WORK

This project, Depressive and Non-Depressive Tweet Classification Using Deep Learning, demonstrates how artificial intelligence can be applied to detect signs of depression through social media content. By utilizing machine learning models such as Support Vector Machine (SVM), Logistic Regression, Random Forest, and XGBoost, the system classifies tweets based on emotional tone and sentiment. The implementation followed a systematic pipeline involving tweet collection, preprocessing, feature engineering, model training, and performance evaluation. The results confirmed the feasibility of identifying depressive expressions in textual data, making this approach valuable for mental health analysis.

In future, this system can be extended by integrating real-time tweet monitoring using Twitter's API to provide continuous analysis. Incorporating deep learning-based NLP models like RoBERTa could further enhance the model's understanding of context and emotion. Additionally, expanding the dataset to include multiple languages and dialects will make the system more inclusive. A practical direction includes developing a mobile or web interface for psychologists or counselors to use this tool in real-world scenarios. Collaboration with mental health professionals can also help in refining the model to ensure ethical and responsible usage. Ultimately, with further development, this project could support early identification of mental health concerns and aid in timely intervention.

REFERENCES

1. A. Orabi, P. Buddhitha, M. Orabi, and D. Inkpen, "Deep Learning for Depression Detection of Twitter Users," Proceedings of the Fifth Workshop on Computational Linguistics and Clinical Psychology, pp. 88–97, 2018.
2. W. Shahzad, M. O. Nusrat, and S. A. Jamal, "Multi-Class Depression Detection Through Tweets Using Artificial Intelligence," International Journal of Advanced Computer Science and Applications, vol. 14, no. 2, pp. 56–65, 2023.
3. A. Cosma, A. M. Bucur, P. Rosso, and L. P. Dinu, "It's Just a Matter of Time: Detecting Depression with Time-Enriched Multimodal Transformers," IEEE Transactions on Affective Computing, vol. 14, no. 1, pp. 215–228, 2023.
4. P. Adibi, M. S. Mahdavinejad, A. Monadjemi, and P. Hitzler, "They Look Like Each Other: Explainable Depression Detection Using Large Language Models," Expert Systems with Applications, vol. 224, 119889, 2023.
5. E. Cambria, S. Han, and R. Mao, "Hierarchical Attention Network for Explainable Depression Detection on Twitter Aided by Metaphor Concept Mappings," Information Fusion, vol. 89, pp. 325–336, 2022.
6. J. Devlin, M. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding," Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics, pp. 4171–4186, 2019.
7. Y. Liu et al., "RoBERTa: A Robustly Optimized BERT Pretraining Approach," arXiv preprint arXiv:1907.11692, 2019.
8. M. Young, The Technical Writer's Handbook, Mill Valley, CA: University Science, 1989.

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

INNO SPACE SJIF Scientific Journal Impact Factor

doi crossref

निस्केयर NISCAIR

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

9940 572 462 6381 907 438 ijircce@gmail.com

www.ijircce.com

Scan to save the contact details