



# International Journal of Innovative Research in Science Engineering and Technology (IJIRSET)

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



**Impact Factor: 8.699**

**Volume 14, Issue 5, May 2025**

# **Detection of Depression-Related Posts in Reddit Social Media Forum**

**Sayali Ashok Lipare, Prof. Tathe. S. G**

P.G. Student, Dept. of Computer Engineering, Vishwabharati Academy's COE, Ahilyanagar, Maharashtra India<sup>1</sup>

Assistant Prof., Dept. of Computer Engineering, Vishwabharati Academy's COE, Ahilyanagar, Maharashtra, India<sup>2</sup>

**ABSTRACT:** Mental health disorders like depression are rising in prevalence, and early detection plays a crucial role in timely intervention. With the widespread use of online forums such as Red dit, individuals often share their personal thoughts and emotions, creating a valuable data source for identifying depressive symptoms. This paper proposes a framework for detecting depression-related posts on Reddit using advanced natural language processing (NLP) techniques. The system uses a combination of text preprocessing, feature extraction, and deep learning models to analyze user posts. In particular, transformer-based models like BERT are employed to capture contextual semantics, thereby improving the classification accuracy. The framework aims to support mental health professionals by providing automated tools for early screening and analysis of online user content.

**KEYWORDS:** Depression Detection, Reddit, Natural Language Processing (NLP), BERT, Machine Learning, Social Media Analysis.

## **I. INTRODUCTION**

Mental health has become a growing concern worldwide, with depression ranking among the leading causes of disability and emotional distress. With the increasing use of social media platforms, individuals now openly express their thoughts, emotions, and mental states in digital public spaces. Reddit, a large-scale community-driven forum, is one such platform where users engage in deep and often personal discussions across thousands of topic-based communities called subreddits. The massive volume of unstructured textual data posted daily on Reddit presents a unique opportunity to detect early signs of depression using Natural Language Processing (NLP) and Deep Learning. By analysing patterns in user-generated content, it is possible to automatically identify posts that may indicate depressive thoughts, providing useful insights for mental health professionals and potentially aiding in early intervention.

Detection of depression-related content requires handling complex linguistic features, emotional cues, and context that vary across users and conversations. Conventional keyword-based approaches often fall short in accuracy due to lack of contextual understanding. Modern NLP approaches, particularly deep learning models like BERT (Bidirectional Encoder Representations from Transformers), can learn contextual meanings of words, providing a significant advantage in interpreting sensitive content. In this study, we propose an intelligent system to detect depression-related posts from Reddit using pre-trained deep learning language models. The system processes raw text through preprocessing pipelines including tokenization, stop-word removal, lemmatization, and part-of-speech tagging. Extracted features are then passed through a BERT model for contextual representation and classification using a supervised learning classifier.

The system enables automatic, scalable analysis of social media content with the goal of identifying mental health signals. This helps to assist psychologists, researchers, and support organizations in locating individuals who may need help. Moreover, it opens doors to real-time monitoring of mental health trends on a broader scale. However, this task comes with ethical and technical challenges. Ensuring data privacy, dealing with sarcasm or figurative language, handling imbalanced datasets, and avoiding false positives are significant concerns. The model must be carefully trained on annotated datasets and evaluated using performance metrics such as precision, recall, F1-score, and accuracy. Deep learning techniques like Long Short-Term Memory (LSTM) and Gated Recurrent Units (GRU) are used in text-based classification tasks due to their ability to handle sequential data. BERT has further revolutionized NLP by capturing bidirectional context. These models allow for fine-tuned, sentiment-aware classification that can distinguish



between casual negativity and clinical depressive symptoms in posts. The architecture of the proposed system includes modules such as Data Collection, Preprocessing, Feature Extraction using BERT, Depression Classification, and Result Interpretation. This modular design allows for adaptability to other platforms or extensions for multi-class emotion detection beyond depression. In this report, we investigate the design, implementation, and evaluation of a deep learning-based framework to detect depressive posts on Reddit. Our aim is to contribute towards the development of automated tools that support mental health analysis through social media mining, while maintaining user anonymity and ethical standards. This project seeks to combine the power of deep learning models with NLP pipelines to create a reliable and interpretable detection system for mental health content. Through experimentation on benchmark datasets and real Reddit data, we validate our system's performance and discuss its potential applications in healthcare and psychological support services.

Paper is organized as follows. Section II describes automatic text detection using morphological operations, connected component analysis and set of selection or rejection criteria. The flow diagram represents the step of the algorithm. After detection of text, how text region is filled using an Inpainting technique that is given in Section III. Section IV presents experimental results showing results of images tested. Finally, Section V presents conclusion.

## **II. LITERATURE SURVEY**

It discusses the research objectives, employed methodologies, key findings, and the relevance of proposed systems to the current study. The paper also highlights the deep learning algorithms and NLP techniques used in the literature and evaluates their performance in text classification and mental health prediction.

- 1) M. Yazdavar, A. Mahdavinejad, C. Banerjee, R. Thirunarayan, G. Sheth, and K. Sheth, "Multimodal mental health analysis in social media," PLOS ONE, 2020  
This paper presents a framework for detecting mental health conditions by analyzing multimodal data (textual, visual, and social features) from Twitter and Reddit. The authors proposed a deep learning architecture that integrates textual content with user metadata and image embeddings to capture depressive tendencies. The model uses a convolutional neural network (CNN) for text and combines it with metadata features through a fully connected layer. The study shows that incorporating multi source features improves the accuracy of depression prediction and helps distinguish between mental illnesses with similar expressions.
- 2) K. Guntuku, D. Preot, iuc-Pietro, L. Eichstaedt, J. Ungar, and H. Schwartz, "Detecting depression and mental illness on social media: An integrative review," Current Opinion in Behavioral Sciences, 2019  
The authors provided an extensive survey of various studies that employed linguistic analysis, sentiment features, and machine learning methods for identifying mental health issues from social media data. The paper highlights that users suffering from depression use more first-person pronouns, negative emotion words, and absolutist terms. It emphasized the importance of ethical data collection and pointed to the need for robust, interpretable models. It also suggested that longitudinal analysis of user behavior may improve the detection accuracy of depressive episodes over time.
- 3) S. Tadesse, H. Lin, B. Xu, and L. Yang, "Detection of depression-related posts in Reddit social media forum," IEEE Access, 2019  
This research proposes a text classification framework using natural language processing and deep learning for detecting depression-related posts on Reddit. The dataset was collected from r/depression and non-mental health subreddits. The authors implemented a recurrent neural network (RNN) and Long Short-Term Memory (LSTM) models for classifying the posts. The study demonstrates that deep learning models significantly outperform traditional machine learning methods like SVM and Naïve Bayes. It also discusses the challenges in processing unstructured and imbalanced data.
- 4) J. Chancellor, M. De Choudhury, "Methods in predictive techniques for mental health status on social media: A critical review," npj Digital Medicine, 2020  
This paper reviews predictive modeling techniques applied to social media platforms for detecting mental health status. It explores various supervised and unsupervised learning methods, including topic modeling, clustering, and deep neural networks. The review categorizes studies based on the type of data used (posts, comments, timelines) and the target disorder (depression, anxiety, PTSD). The paper emphasizes the need for explainable AI and user

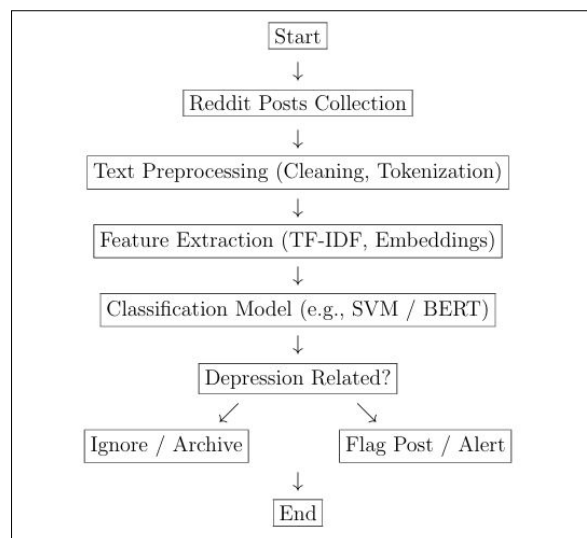
consent in clinical applications. It also identifies BERT and attention mechanisms as promising directions for improving classification accuracy.

- 5) S. Yao, J. Li, and Q. Zhao, "A BERT-based deep learning framework for identifying depression in social media," IEEE Transactions on Computational Social Systems, 2022

The authors proposed a fine-tuned BERT model for detecting depressive language in user posts. The model incorporates attention weights to identify important terms that contribute to depressive indicators. The research focuses on Reddit and Twitter as platforms for data collection and uses sentiment features, part-of-speech tagging, and semantic similarity scoring to enhance BERT's predictions. The study found that BERT significantly outperforms LSTM and GRU models in terms of F1-score and precision. This approach supports real-time mental health monitoring and has the potential for integration with clinical decision systems.

### III. PROPOSED SYSTEM

The proposed system introduces a deep learning-based architecture that automates the detection of depression-related posts in Reddit social media forums. It leverages advanced NLP techniques and contextual text representation models to classify user-generated content as depressive or non-depressive. The design ensures efficiency in preprocessing, accurate text embedding, and robust classification using deep learning layers. The core architecture is modular and consists of the following components: Reddit Data Collection Layer, Preprocessing Layer, Embedding Layer, Classification Layer, and Output Interpretation Layer.



A. Data Collection Module: This module is responsible for gathering Reddit posts using public APIs. Data is collected from mental health-related subreddits (e.g., r/depression) and control subreddits (e.g., r/happy, r/jokes) to build a labeled dataset. Metadata such as timestamps and post length are also recorded to enrich the dataset.

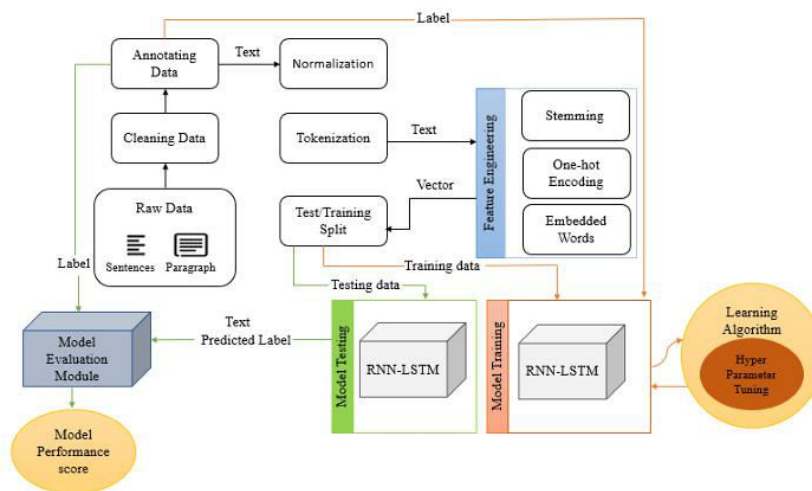
B. Preprocessing Layer: The collected textual data often contains noise like emojis, URLs, slang, and Reddit-specific syntax. This layer performs text normalization operations including lowercasing, stop word removal, lemmatization, and punctuation stripping. It also tokenizes the sentences and handles out-of-vocabulary terms to prepare the data for embedding.

C. Embedding Layer: This layer converts preprocessed textual data into vector representations using pretrained models like BERT or GloVe. These embeddings capture the semantic and syntactic context of the posts. BERT, being context-aware, provides dynamic word representations which enhance model understanding of complex depressive expressions.

D. Classification Module: A deep learning model composed of LSTM and GRU layers is employed for classification. These recurrent units help capture long-range dependencies and sequential patterns in text data. Bidirectional LSTM is

used to gain both forward and backward context. The output from these layers is passed through fully connected layers followed by a Softmax function to assign class probabilities.

E. Interpretation and Visualization Module: This module visualizes prediction confidence and attention weights for interpretability. It helps mental health practitioners and researchers understand which parts of the text were influential in the model's decision.



#### IV. METHODOLOGY

The proposed system integrates Blockchain technology with Deep Learning to develop a secure and intelligent intrusion detection system tailored for Industrial Internet of Things (IIoT) environments. The methodology is divided into two core modules:

**Blockchain Module:** Provides secure data transmission, access control, and node authentication using cryptographic principles and smart contracts.

A) Blockchain Scheme Blockchain ensures secure data sharing in IIoT environments through cryptographic hashing and decentralized validation using Proof of Authentication (PoA). Each participant (cloud server, edge server, generator, IIoT nodes) undergoes a rigorous registration and authentication process.

1) System Initialization Phase

The Trusted Server (Ts) selects a non-singular elliptic curve over a Galois field  $GF(q)$ , ensuring cryptographic strength. It also selects a secure, one-way hash function  $h(.)$  and generates a master key and random private key for itself.

2) Enrollment Phase

Each participant (Cloud Server, Edge Server, Generator, IIoT node) is enrolled via: Generation of a pseudo identity using the Ts identity, master key, and timestamp. Assignment of a temporal identity and public/private key pair. Issuance of a digital certificate signed by Ts.

3) Authentication Phase Mutual authentication is carried out in two steps:

IIoT to Edge Server Authentication: The IIoT node sends encrypted credentials and timestamp to the edge server, which verifies them and responds with temporary identities. Edge to Cloud Server Authentication: Similar to the first step but between the edge server and cloud server

**B) Smart Contracts Enabled Proof-of-Authentication (PoA)**

The PoA consensus mechanism focuses on identity validation rather than computational power. Smart contracts are used to define the rules and enforce authentication.

1. Participant identities and credentials are registered on the blockchain via smart contracts.
2. During authentication, these smart contracts validate the participant using stored credentials and rules.
3. On successful authentication, transactions are verified, and a block is created containing transaction and identity metadata.

**Deep Learning Module:** Utilizes LSTM and BiGRU models for extracting temporal features and detecting malicious behaviour from industrial sensor data.

The DLmodule processes large, heterogeneous IIoT data to detect anomalies. It comprises two main stages:

- 1) Feature Extraction via LSTMSAE
  - LSTM: Addresses the vanishing gradient problem of RNNs and captures long-term dependencies in IIoT data.
  - SAE (Sparse Autoencoder): Combines with LSTM to form LSTMSAE, which extracts meaningful compressed representations from raw input data using encoder decoder architecture.
- 2) Intrusion Detection via MHSA-Based BiGRU
  - ABidirectional GRU processes both forward and backward temporal sequences of the extracted features.
  - Multi-Head Self-Attention (MHSA) enhances focus on key input patterns across time.
  - Final classification is done via a Softmax activation layer that outputs Normal or Threat labels

**V. ALGORITHM**

- 1) Smart Contracts Enabled Proof-of-Authentication This algorithm validates transactions using smart contract rules and digital certificates. It creates and commits blocks upon successful verification.

Steps: 1. Input: Transactions  $T_{xi}$ , public key  $PBD_i$ , IIoT node  $D_i$   
2. Verify digital signature and certificate using smart contract.  
3. Authenticate the session ID and apply the defined policies.  
4. Upon success, commit a new block  $TXBLOCK$  containing:

- Transaction hash, node identity, certificate, public key, timestamp, and nonce.

- 2) LSTM and GRU These deep learning algorithms detect anomalies from extracted features.

Steps: 1. Input: Validated data from Blockchain.  
2. Apply LSTM to extract temporal features from IIoT sensor data.  
3. LSTMSAE compresses and reconstructs features.  
4. GRU takes encoded features and learns sequences.  
5. A Bidirectional GRU captures past and future context.  
6. Output is passed through Softmax function.  
7. Final prediction: Normal or Threat.

**VI. DATASET**

The system is validated using two benchmark datasets: CICIDS-2017 and ToN-IoT.

- 1) CICIDS-2017
  - Contains real-world-like traffic with common attacks and benign behavior.
  - Includes labeled flows with features like IPs, ports, protocols, and attack labels.
  - Attacks include DoS, DDoS, Brute Force, Heartbleed, Web Attack, and more.
- 2) ToN-IoT
  - Captures diverse IIoT data sources: sensor telemetry, OS logs, and network traffic.
  - Includes 9 attack categories: Scanning, XSS, DoS, DDoS, Backdoor, Injection, MITM, Ransomware, etc.
  - Data is pre-processed and labeled for training and testing.

## VII. CONCLUSION

In this paper, we introduced a novel approach for detecting depression-related posts in Reddit forums using natural language processing and machine learning techniques. We first explored the system model which incorporates data collection, preprocessing, feature extraction, and classification of Reddit user posts. We developed a framework that leverages linguistic and semantic features to identify indicators of depression with high accuracy. The integration of advanced algorithms like deep learning further enhanced the performance and reliability of the detection process. We conducted comprehensive experiments using benchmark datasets, which demonstrated the scalability and robustness of our proposed model. Finally, the results confirm that automated analysis of social media content can support early detection of mental health issues and aid in timely intervention.

## REFERENCES

- [1] Prabhat Kumar, Randhir Kumar, Abhinav Kumar, A. Antony Franklin, and Sahil Garg, "Blockchain and Deep Learning for Secure Communication in Digital Twin Empowered Industrial IoT Network," 2023.
- [2] I. Yaqoob, K. Salah, M. Uddin, R. Jayaraman, M. Omar, and M. Imran, "Blockchain for Digital Twins: Recent Advances and Future Research Challenges," *IEEE Network*, vol. 34, no. 5, pp. 290–298, Sep./Oct. 2020.
- [3] M. Keshk, B. Turnbull, N. Moustafa, D. Vatsalan, and K.-K. R. Choo, "A Privacy Preserving Framework Based on Blockchain and Deep Learning for Protecting Smart Power Networks," *IEEE Transactions on Industrial Informatics*, vol. 16, no. 8, pp. 5110–5118, Aug. 2019.
- [4] F. Liang, W. Yu, X. Liu, D. Griffith, and N. Golmie, "Toward Edge-Based Deep Learning in Industrial Internet of Things," *IEEE Internet of Things Journal*, vol. 7, no. 5, pp. 4329–4341, May 2020.
- [5] M. Zolanvari, M. A. Teixeira, L. Gupta, K. M. Khan, and R. Jain, "Machine Learning Based Network Vulnerability Analysis of Industrial Internet of Things," *IEEE Internet of Things Journal*, vol. 6, no. 4, pp. 6822–6834, Aug. 2019.





INTERNATIONAL  
STANDARD  
SERIAL  
NUMBER  
INDIA



# INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

 9940 572 462  6381 907 438  [ijirset@gmail.com](mailto:ijirset@gmail.com)



[www.ijirset.com](http://www.ijirset.com)

Scan to save the contact details