



International Journal of Advanced Research in Education and Technology (IJARETY)

Volume 13, Issue 3, May - June 2026

Impact Factor: 9.971



Mental Health Prediction Using Journal Entries

Geetha L

MCA Student, CMR Institute of Technology, Bangalore, India

ABSTRACT: Research focused on understanding mental health using text-based data is expanding globally because emotional well-being influences individuals from various age groups. The early recognition of conditions such as depression, anxiety, and emotional strain can motivate Encourages individuals to obtain help at the right time and supports improved mental health outcomes. Personal diary entries document individual thoughts, emotions, and experiences, making them a meaningful source for mental health evaluation. several investigators have explored how documented content can assist uncover patterns connected to emotional conditions. Strategies comprising opinion analysis, language processing techniques, and machine learning are often utilized for this purpose. Different approaches have been developed, but they remain face challenges to determine deeper emotions from unorganized text. To overcome this problem, the paper develops a framework for assessing emotional condition from personal journals. The system analyzes textual data using language processing and transformer-based feature extraction to gather meaningful data from the text collected characteristics are subsequently applied to recognize the label using a ML model that can group different types of emotional health conditions.

KEYWORDS: Mental health prediction, Journal entries, Natural language processing, Machine learning, Transformer models.

I. INTRODUCTION

Mental health is an general well-being, as it clearly shows how individuals think, experience emotions, and act in their day-to-day activities. A balanced mental state allows people to deal with stress, sustain relationships, and handle routine activities effectively. When mental health conditions such as depression, anxiety, and emotional stress remain unrecognized or untreated, they can reduce quality of life and negatively affect social and professional functioning [1]. Therefore, timely detection and continuous monitoring of mental health issues is important for timely support and long-term wellbeing. Textual data is an important origin for understanding mental health conditions, as written language often reflects emotional and psychological states. Past research has revealed that sentiment analysis, natural language processing, and machine learning techniques can effectively detect depression and emotional distress from text data [1], [4]. Natural language approaches support clean, structure, and analyze unstructured text by extracting meaningful linguistic and emotional patterns. Compared to traditional Beyond standard clinical evaluations, text-based mental health analysis enables flexible and unobtrusive mental health monitoring in real-life scenarios. Current studies have shown increasing interest in transformer-based learningbased models for predicting mental health conditions due to their strong ability to capture contextual and emotional information. BERT-based approaches have demonstrated improved performance in depression detection tasks using social media text [2]. Hybrid transformer architectures and contextualized embeddings have further enhanced enhanced prediction results using modelling complex semantic and emotional patterns [3], [5]. In addition, machine learning and transfer learning strategies have been implemented applied to detect multiple mental health conditions from online platforms such as Reddit [4]. The approach unites language processing frameworks that make use of ML technique to explore textual information overcome the constraints found in traditional approaches, this study develops a methodology that recognizes mental well-being challenges from personal narrative entries. personal diary entries are chosen as the key information source because people often convey their authentic feelings, perspectives, and routine activities in them. Consequently, these entries meaningful indications of an individual's emotional condition. In the designed model, starts with a preprocessing stage where journal entries are processed using textprocessing methods. During this step, the text is cleaned, unnecessary information is removed, and the data is prepared for further analysis. During this phase, the content is improved by eliminating irrelevant elements and organizing the data for the next processing step. When the text preparation is completed transformer models are applied to understand deeper meaning and emotion-related patterns appearing in the entries. collected features is then used train ML classification models capable of recognizing mental health states such as stress, anxiety, and depression. The aim of this system is to identify emotional trends early while protecting privacy safeguarding user data reducing impact into day-to-day activities. The remaining part of this paper is arranged as follow II reviews examines earlier studies related to detecting mental health issues using written data. Section III presents the research method applied in this study, including initial processing steps, feature derivation, and the prediction process.

Section IV presents the experimental results of the system. Section V explains the obtained results their explanation. Finally, Section VI gives the conclusion and recommends areas for future work.

II. PROPOSED SYSTEM / ARCHITECTURE

The proposed system is designed to predict mental health states from personal journal entries using natural language data analysis combined with learning algorithms techniques. Figure 1 illustrates the overall architecture presented system mental health prediction system. The system utilizes a stepby-step framework that processes raw journal text into meaningful predictions representing different mental health level.

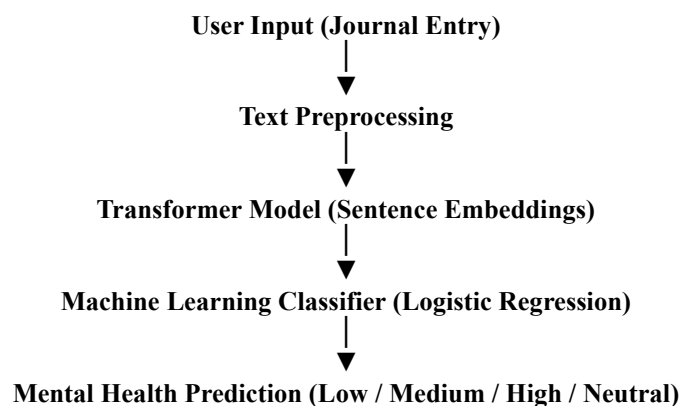


Fig. 1. Proposed System Architecture

A. User Input Module

The process begins when a user enters a journal entry into the system. The input consists of personal thoughts, emotions, and daily experiences that may contain related to mental health conditions. These journal entries serve as the input data for analysis, as they often reflect real emotional states in a private and unbiased manner.

B. Text Cleaning and Preprocessing

The input text is created with the help of natural language processing methods This step includes converting the text is developed through NLP-based methods irrelevant symbols, and normalizing the text. Preprocessing reduces noise and improves the quality of data used for feature extraction. improves the quality of data used for feature extraction.

C. Sentence-BERT Embedding

After preprocessing, the cleaned text is changed into numerical vectors using a Sentence-BERT model. Sentence-BERT generates dense semantic embeddings that capture contextual and emotional information present in the journal text.

D. Logistic Regression Classification

The generated embeddings are then passed to a Logistic Regression classifier trained on mental health–related data. The classifier analyzes the extracted features and estimates the probability of different mental health states.

E. Prediction Output

Finally, the system produces a prediction representing the user’s mental health state, categorized into levels such as Low, Medium, High, or Neutral. This output supports early mental health awareness and assists in identifying emotional patterns in a non-intrusive manner

III. RESULTS AND DISCUSSION

A. Tools Used

The system was put together using a set of commonly used tools and libraries. Python was picked as the main language since it offers good support for data handling, machine learning, and NLP work. Sentence-BERT was used to pull out semantic embeddings from journal text, helping the system pick up on contextual and emotional signals. The Scikit-learn library handled the Logistic Regression part for classifying mental health states. Streamlit was brought in to build

a straightforward web interface where users can type their journal entries and get predictions right away. Putting these tools together gave a working and adaptable setup for building and testing the mental health prediction system.

B. Emotional State Recognition

This produced result demonstrates a general affective situation determination, intensity of tension, nervousness, and sadness. The platform reviews recorded personal entries and determines affective status by recognizing patterns in written content. Meaningful sentence representations are derived using Sentence-BERT, and these features are then evaluated through a Logistic Regression model to determine emotion levels. Every feeling category is labeled with a level of Low, Medium, or High according to the noticed emotional behavior.

User Input: "I feel lonely and nobody understands me."

Predicted Output:

- Stress: Medium
- Anxiety: Low
- Depression: High
- Overall Mood: Negative

This result shows that the system correctly picks up on emotional cues like loneliness and isolation, which are commonly linked to depressive states. It shows the model can tell apart different emotional conditions without raising unnecessary alerts. This kind of output makes early emotional awareness more accessible in a simple, user-friendly way.

C. Aggression State Recognition

When a user input contains language indicating potential harm to others, the system detects the content as **aggressive** and displays a warning message.

Example Input: "I feel like hitting my friend."

System Output: Aggressive or violent language has been recognized. If you think you may hurt someone, stop and seek help right away.

In this scenario, the system **does not treat the input as a mental health condition**. Instead, it interprets the response as **stressrelated aggressive behavior** and issues a safety alert. This functionality demonstrates the system's capacity to detect rising emotional intensity and promote **selfcontrol and preventive action** before harmful behavior occurs.

D. Crisis State Detection

The third output type represents crisis situations involving self-harm or suicidal intent. When such language is detected, the system immediately prioritizes safety and generates an emergency alert.

Example Input: "I am more stressed and want to die."

Crisis Detected — Immediate Assistance Required

The input text indicates possible self-harm or suicidal intent. Please seek immediate help by contacting emergency services or a trusted person.

Helplines (India): • AASRA: +91-9820466726 •

Snehi: +91-9582208181 • Sumaitri: 011-23389090

Other countries: findahelpline.com

You are not alone. Immediate support is available.

When a crisis-related entry comes in, the system skips the emotion classification step and instead brings up helpline and emergency contact details directly. This shows the system treats user safety as the top concern, especially when the situation is serious.

IV. CONCLUSION

This work presents a text-based approach to support mental health awareness by looking at what people write in their personal journals. The system uses NLP and machine learning to go through written entries and figure out emotional states like stress, anxiety, and depression. It starts with cleaning the text, then builds sentence-level representations with SentenceBERT, and finally runs them through a Logistic Regression model to gauge emotional levels.

Safety checks are also built in to catch any language around aggression or crisis situations. When that kind of content appears, the system puts out a warning or emergency support details instead of carrying on with normal emotional

analysis. Since the system works with private journal entries rather than public social media posts, it offers a quieter and less intrusive way to keep track of mental health.

The test outputs show that the model can pick up on emotional shifts and respond in a fitting way. On the whole, this approach points to the value of using journal-based text analysis as a practical support tool for early mental health monitoring and awareness.

REFERENCES

- [1] S. Zhou and M. Mohd, "Mental Health Safety and Depression Detection in Social Media Text Data: A Classification Approach Based on a Deep Learning Model," IEEE Access, vol. 13, 2025.
- [2] A. Raj, Z. Ali, S. Chaudhary, K. K. Bali, and A. Sharma, "Depression Detection Using BERT on Social Media Platforms," in 2024 IEEE International Conference on Artificial Intelligence in Engineering and Technology (IICAIET), 2024, IEEE.
- [3] A. Karamat, M. Imran, M. U. Yaseen, R. Bukhsh, S. Aslam, and N. Ashraf, "A Hybrid Transformer Architecture for Multiclass Mental Illness Prediction Using Social Media Text," IEEE Access, 2025.
- [4] A. Takdenti and A. C. Mazari, "Multi Aspect Detecting Mental Disorders Using Machine and Transfer Learning Models on Reddit Posts," in 2025 International Conference on Intelligent Computer Systems, Data Science and Applications (IC2SDA), 2025.
- [5] M. A. Abbas, K. Munir, A. Raza, N. A. Samee, M. M. Jamjoom, and Z. Ullah, "Novel Transformer Based Contextualized Embedding and Probabilistic Features for Depression Detection From Social Media," IEEE Access, vol. 12, 2024.
- [6] N. Smetanin, "The Applications of Transfer Learning to Mental Health Classification from Social Media Text," Frontiers in Psychiatry, vol. 11, 2020.
- [7] S. Agrawal and A. An, "Unsupervised Emotion Detection from Text Using Semantic and Syntactic Relations," in Proc. IEEE/WIC/ACM International Conference on Web Intelligence, 2012.
- [8] R. Lopes, N. Rodrigues, and F. Salgado, "Emotion Recognition from Text Using Sentence-Level Embeddings and Attention Mechanisms," Expert Systems with Applications, vol. 210, 2023.
- [9] K. Cobbe, V. Kosaraju, M. Bavarian, M. Chen, H. W. Choi, R. Child, S. Schulman, I. Sutskever, and J. Ziegler, "Mental Health Monitoring via Personal Diaries: An NLP Perspective," Journal of Biomedical Informatics, vol. 118, 2021.
- [10] Y. Kim, "Convolutional Neural Networks for Sentence Classification," in Proc. Conference on Empirical Methods in Natural Language Processing (EMNLP), 2014.



International Journal of Advanced Research in Education and Technology

ISSN: 2394-2975

Impact Factor: 9.971