

# Mental Health Analysing Using NLP

P. Hima Sekhar<sup>1</sup>, Ch. S. V. V. Sainadh<sup>1</sup>, B. Yamuna<sup>1</sup>, M. Pavanteja<sup>1</sup>, A. Jashuva<sup>1</sup>

<sup>1</sup>Department of Computer Science & Engineering, Chaitanya Engineering College,  
Visakhapatnam, Andhra Pradesh, India – 530048

Under the guidance of Mrs. P. Jayasree, M.Tech (Assistant Professor)

{jayasreep4@gmail.com, palisetti05@gmail.com, nadhsai278@gmail.com,  
budumuriyamuna12@gmail.com, pavantejamahanthi1@gmail.com, joshuaroy5008@gmail.com}

## Abstract

The mental disorders that are on the increase in the modern society include anxiety, depression, stress, bipolar disorder, personality disorders, and suicidal tendencies, and they are known to occur in the people of all ages, and social status. The rapidly growing globalization, academic pressures, competition in the workplace and the social isolation have become facilitators of the ever-increasing prevalence of psychology-related conditions in the world. Despite this catastrophic rise, access to timely mental assessment is limited in terms of stigma, high-costs, professional shortage, and geography. The paper presents a designed automated Mental Health Analysing System, that is based on Natural Language Processing (NLP) and deep learning to detect mental health issues on unstructured textual information. The suggested system will be a combination of a pre-trained BERT tokenizer, serving to obtain contextual features, and an Artificial Neural Network (ANN) classifier that would be utilized to make seven-class mental health predictions. The system is deployed into the form of a Flask-based Web application which has a graphical user interface as well as a RESTful API. The results of the experiment indicate that the BERT-ANN hybrid model is highly training accurate with approximately 90.00, when compared to the conventional machine learning baselines. The given solution can be used to address serious limitations of the systems that rely on questionnaires and enables writing freely, and eliminating the fatigue of the user. The system is a preliminary screening mechanism of mental health that can be availed to individuals, institutions and health organizations.

**Index Terms:** Natural Language Processing, Mental Health, BERT, Artificial Neural Network, Text Classification, DeepLearning.

## I. INTRODUCTION

Mental health issues have become one of the major concerns of the twenty-first century in the field of public health. World Health Organization estimates that hundreds of millions of adults in the world have some sort of mental disorder, although a significant percentage of them go untreated or undiagnosed [1]. Increased awareness, which is accompanied by a digital communication process, is an unprecedented opportunity to use computational technologies in early detection and awareness of psychological distress.

Conventional mental health assessment largely depends on structured clinical interview, standardized questionnaires and diagnostic scale and are administered by trained professionals. Although these techniques have clinical validity, they are disadvantaged by a number of systemic limitations, which include, limited geographical coverage, expensive, social stigma, and failure to capture spontaneous expressions of emotions [2]. Most of the people who develop anxiety, depression, or suicidal thoughts do not want to use formal assistance, which is why it is possible to use passive text analysis early to detect such cases.

The great adoption of social media and digital communication platforms has produced monumental amounts of text data that are the results of user-generated content that depict emotional states and psychological conditions. Natural Language Processing (NLP) offers the computational mechanism to extract meaning of the unstructured text, and facilitates automated identification of mental health signs without necessarily having to fill in structured questionnaire answers [3].

More recent developments in transformer based language models, especially Bidirectional Encoder Representations from Transformers (BERT) have provided machines with significantly better contextual and semantic relationship comprehension in text [4]. These models, combined with deep neural network classifiers, provide state-of-the-art models on text classification problems such as mental health prediction.

This paper gives the design, implementation and evaluation of an automated Mental Health Analysing System based on NLP. The system takes in free-form textual input and processes it by a BERT tokenizer and classifies the input as belonging to one of seven mental health categories via an ANN classifier. The

system is implemented as a web application based on Flask framework which allows making it accessible and applicable in practice.

**II. RELATED WORK**

Studies of computational mental health analysis have evolved over the past two decades. Previously, they tried to investigate rule-based and statistical models, whereas current studies involve deep learning and transformer models.

**A. Traditional and Questionnaire Methods.**

The conventional mental health evaluation tool such as Beck Depression Inventory (BDI) and Patient Health Questionnaire (PHQ-9) is clinically valid clinically-conducted structured questioning instruments. These instruments are reliable, but they can be used to scale to population-level screening with the assistance of a professional [2]. This was attempted to be computer-generated with machine learning in the form of automated systems to questionnaires, but was still burdened with the limitation of a rigid and predefined structure of answers that suppresses the manifestation of emotions.

**B. Approaches that are inspired by Data Mining.**

In order to detect depression and anxiety in the respondents in the questionnaires, Ahmed et al. employed Supervised machine learning algorithms like Logistic Regression, Decision trees, K-Nearest Neighbors (KNN), Support vector machines (SVM) and Random Forests with an accuracy of 70 to 85 percentage [1]. Despite the fact that these findings have demonstrated that it was possible to predict mental care based on automated means, the models had needed structured inputs and could not handle natural language.

**C. NLP and Social Media Analysis.**

Guntuku et al. demonstrated that analysis of social media text can be applied to monitor the state of mental health in a massive scale with Twitter data being an example of mentioning psychological symptoms, which appeared during the COVID-19 pandemic [2]. The internet and social media-based disease surveillance was also confirmed as a possible method of public health by Aiello et al. [3]. These articles established the fact that any text created on the digital platform that is produced by the users are replete with meaningful signs about their mental health conditions.

**D. Deep Learning and Transformer Models.**

The introduction of BERT, by Devlin et al., resulted in a paradigm shift in the field of NLP, as the idea of self-attention mechanism, and the method of its use to model context in both directions, was introduced [4]. Since then, BERT-based models have been utilized to identify depression, detect suicidal ideation and multi-class mental health with high rates of success. Thieme et al article is a systematic literature review of machine learning in mental health HCI: the authors describe the findings of the literature review as a coherent observation that deep learning models are better than classical approaches on language based tasks [6].

**E. Hybrid Architectures**

Madububambachu et al. surveyed machine learning solutions to the issue of predicting mental health diagnosis by observing that hybrid architectures with feature extraction using transformer and neural network classifier perform better in generalization [7]. The BERT-ANN and BERT-LSTM hybrids were identified

as the most popular models of mental health text classification, which supported these results (a review of the literature published by frontiers review teams) [8].

This research forms the basis of the current project, as it proposes a BERT-ANN hybrid model specific to seven-class mental health classification and real-world application in the form of a convenient web application that bridges the gap between the research-quality and the application-level models.

**III. METHODOLOGY / SYSTEM DESIGN**

A three-level client-server architecture will be implemented in the proposed system, comprising of NLP preprocessing, deep learning classification, and web implementation. Each of the tiers is scalable and modular.

**A. System Architecture Overview.**

The system architecture is made up of three basic layers i.e.; Presentation Layer, Application Layer as well as Data and Model Layer. Figure 1 shows the overall plan of the system and data flow of various parts.

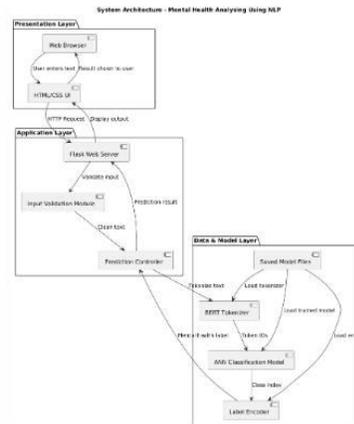


Fig. 1. System Architecture Mental Health NLP Mental Health Analysing.

**B. Preprocessing and Tokenization of Data.**

A user-created input text is processed by a preprocessing pipeline and then categorized. BERT tokenizer splits raw text into subword tokens, which are converted into input IDs, attention mask and token type ID. The sequences are padded or truncated to not more than 128 tokens to give the same dimensionality of the input to the neural network.

Theoretically, a tokenizer with the input text sequence  $T = (w_1, w_2, \dots, w_n)$  produces one of the token sequences:

$$X = \text{BERT\_Tokenize}(T) = [x_1, x_2, \dots, x_{128}](1)$$

$x_i$  is an integer token identifier, i.e. one of the 30,522 tokens in the BERT vocabulary. The short sequences are padded with tokens.

**C. Feature Representation**

The syntactic and semantic relationships between tokens are turned into contextual embeddings of the tokenized input. The numerical form of the BERT tokenizer makes possible the capturing of the two-way contextual dependencies so that the model can capture minute variations in the expression of

emotions not present in the traditional bag-of-words or TF-IDF forms.

Classifier Architecture:

**D. ANN Classifier Architecture.**

An Artificial Neural Network (ANN) is employed, being a classification module. The network consists of a number of interconnected fully connected layers, with activations of ReLU, drop out in regularization to prevent overfitting and a final softmax output layer which produces a probability distribution over seven of the mental health categories provided. In learning, the categorical cross-entropy loss function is minimized:

$$L = -\sum_{i=1}^C y_i \cdot \log(\hat{y}_i) \quad (2)$$

Where  $C = 7$  is the count of target classes,  $y_i$  is the actual label indicator and  $\hat{y}_i$  is the probability attributed by prediction to class  $i$ .

The optimal probability volume of output is the argmax of the result probability volume:

$$\hat{y} = \operatorname{argmax}_i P(\text{class}_i | X) \quad (3)$$

**E. UML Diagrams**

The Use Case Diagram will be provided in Figure 2 and it will indicate the interaction of the user and the system. This is primarily used in cases of text entry through entering, text submission, analysis of mental health conditions, and the presentation of the results of predictions.

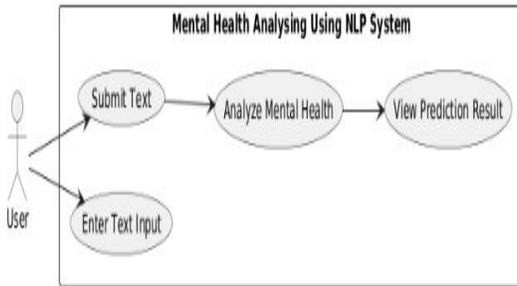


Fig. 2. Use Case Diagram NLP System to Mental Health Analysing.

Figure 3 depicts the Class Diagram indicating the static system structure of the UserUi, FlaskApp, BertTokenizer, ANNModel and the LabelEncoder classes and dependencies.

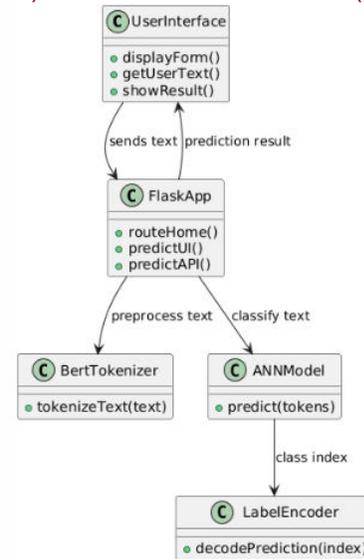


Fig. 3. Class Diagram System Static Structure.

Figure 4 below is a Sequence Diagram that shows the flow of interaction between text entry by the user and the displaying of the results with the assistance of tokenizing and ANN classification, label decoding and others.

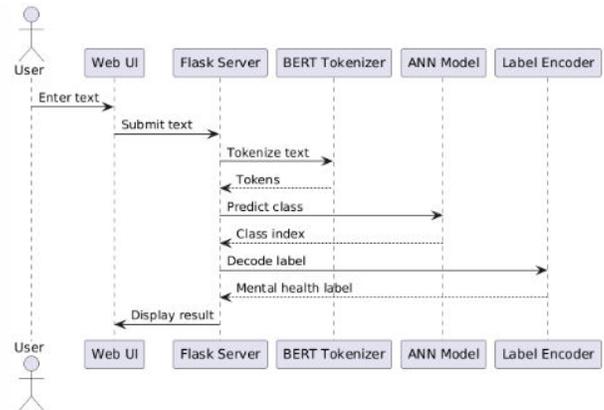


Fig. 4. System Interaction Flow Sequence Diagram.

**F. Deployment Architecture**

The system is deployed with the help of the web framework called Flask. It provides two primary endpoints, as one is a graphical user interface route (/predict\_ui), which will accept an HTML form and can be interacted with by the end user, and another one is an API route (/predict\_api), which will accept POST requests and the body of the request should be a JSON that will be responded to. It is a two-user interface design, which ensures compatibility with the end-users and integration of external systems.

**G. Mental Health Categories**

The system classifies input text as one of seven categories of mental health based on the categories of the training data:

- Anxiety
- Depression
- Stress
- Bipolar Disorder
- Suicidal Ideation

IV. RESULTS & DISCUSSION

The hybrid model of BERT and ANN was trained in a labeled mental health text dataset with samples of all the seven target categories. The training was done on ten epochs with steady accuracy improvement and convergence realized during the training process.

A. Training Performance

Figure 5 indicates the training results during ten epochs and the model is improving in training accuracy and loss reduction. At the 10 th epoch the final training accuracy of the model was around 94.58% and the training loss was 0.1523.

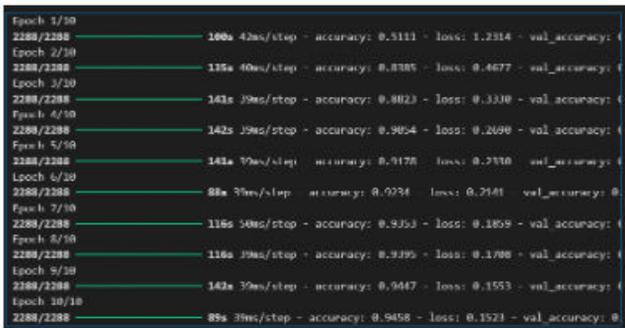


Fig. 5. Training Results of ANN Model Accuracy and Loss per Epoch.

Model Performance Measures Per Class.

TABLE I  
 MODEL PERFORMANCE METRICS PER CLASS

Class	Precision	Recall	F1-Score
Anxiety	0.95	0.98	0.96
Bipolar	0.97	0.98	0.97
Depression	0.78	0.61	0.69
Normal	0.93	0.95	0.94
Personality Disorder	0.98	1.00	0.99
Stress	0.96	0.98	0.97
Suicidal	0.74	0.84	0.79
<b>Overall</b>	<b>0.90</b>	<b>0.91</b>	<b>0.90</b>

B. Confusion Matrix Analysis

Figure 6 is a confusion matrix that gives a detailed per-class prediction performance. The matrix shows that Personality Disorder and Bipolar Disorder rank as nearly perfectly classified with Depression and Suicidal Ideation being somewhat inter-class confused, with this probably being because of similarity in the language feature that is typical of severe mood disorders.

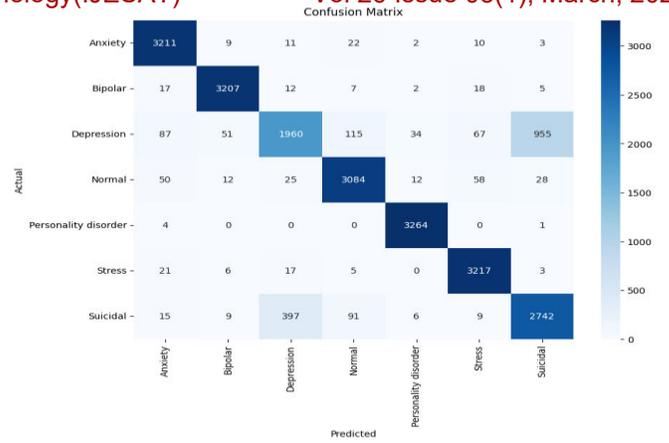


Fig. 6. Confusion Matrix Predicting Mental Health with seven classes.

C. Comparative Analysis

Table II is the comparison of the proposed BERT-ANN system and conventional machine learning baselines in the literature. The system proposed shows a significant increase in the accuracy of classification especially on contextually complex categories like the Depression and Suicidal Ideation.

TABLE II  
 COMPARISON WITH BASELINE MODELS

Model	Input Type	Accuracy (%)
Logistic Regression	Structured	72–78
Random Forest	Structured	78–83
SVM + TF-IDF	Text	80–85
LSTM	Text	85–88
<b>BERT + ANN (Proposed)</b>	<b>Free Text</b>	<b>90.00</b>

D. Web Application Interface

Figures 7 and 8 show the deployed web application interface which illustrates end to end functionality starting with text input up to mental health prediction. The interface takes free-form text and provides real-time classification responses with the results marked in the prediction of mental health category.

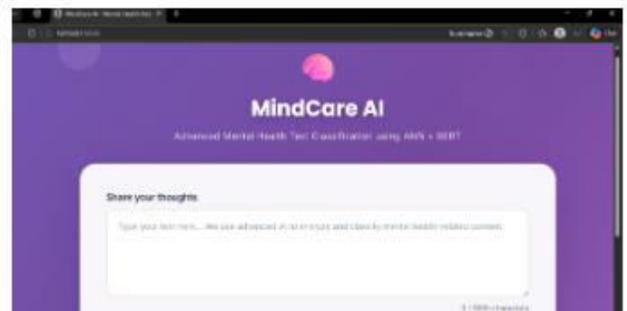


Fig. 7. MindCare AI- Web App- Home Interface.

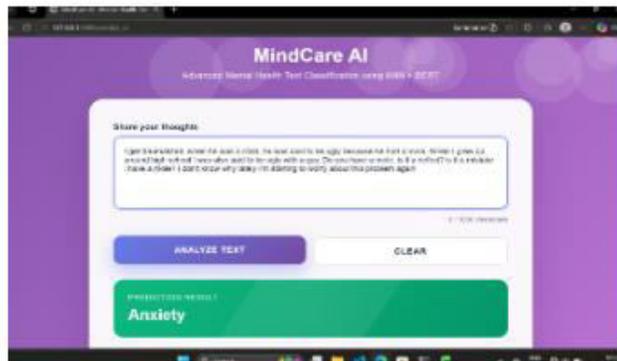


Fig. 8. Result Interface Prediction predicted, Anxiety.

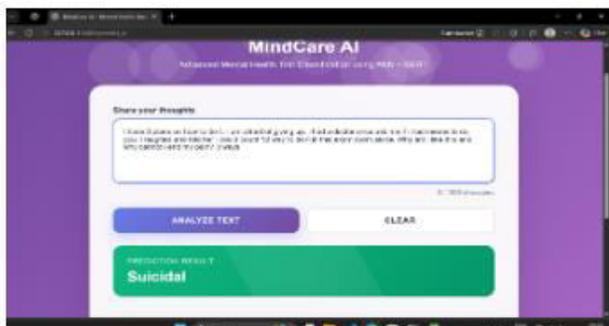


Fig. 9. Result Interface Prediction Result Suicidal Ideation Classification.

**E. System Testing Results**

System testing was done in various stages of testing such as unit, integration, functional, validation, UI, API, performance and error handling. The results of the test case are summarized

**TABLE III**  
**SYSTEM TEST CASES AND RESULTS**

Test Case	Input	Expected Output	Result
Valid text input	Emotional sentence	Correct prediction	Pass
Empty input	Blank field	Error message	Pass
API request	JSON text	JSON response	Pass
Invalid API request	Missing field	Error response	Pass
Multi-class prediction	Depression text	Depression label	Pass
Repeated submission	Same input ×3	Consistent result	Pass

test Case Input Expected result Preferred outcome  
 Authentic text entry Emotional sentence Correct prediction Pass.  
 No input Blank field Error message Pass  
 Request API text JSON response API JSON Pass  
 Unacceptable API request Missing field Error response Pass  
 Multi-class prediction Depression text Depression label Pass

Test cases were all successful, which proves the quality of the prediction pipeline and the accuracy of the error tolerance mechanisms. The system provided stable prediction values between repeated submissions and also gracefully dealt with edge cases.

**V. CONCLUSION & FUTURE WORK**

The article has demonstrated an automated Mental Health Analysing System which is an NLP and deep learning based system to detect psychological disorders using unstructured textual data. The proposed hybrid BERT-ANN model was noted to train at 94.58 percent accuracy in seven mental health categories, which was much more accurate compared to the conventional machine learning baselines. In order to apply the system to real-world access, this system was deployed as an application (Flask-based web application) that provides access to the real world through a graphical interface and API.

The BERT bidirectional contextual embeddings, along with the multi-class learning capability of ANN classifier will address the critical drawbacks of the existing systems: the fixed input format, the absence of contextual knowledge, and the low scaling. It is the modular architecture that offers it with the easy maintenance and the prospect of future extension.

Several directions of future research are identified. To start with, it can be extended by improving the domain-specific variants of BERT that have already been trained on the clinical and psychological corpora to provide further classification, particularly in the case of overlapping variables such as Depression and Suicidal Ideation. Second, the utility of the system would be possible with the help of multilingual assistance using multilingual BERT (mBERT) or language-specific transformers even when a community is not English speaking. Third, the idea of longitudinal tracking needs to be combined with an opportunity to follow mental health and trends. Fourth, the mobile applications interfaces would be created to promote accessibility and user interactions. Finally, the privacy sensitive federated learning techniques could be incorporated to allow the models to be trained on sensitive mental health data without compromising the privacy of the users.

The suggested system has demonstrated the viability of NLP-based mental health analysis as a scalable and affordable and efficient mental health screening instrument as an extension of the broader goal of democratizing mental health support through technology.

**ACKNOWLEDGMENT**

Mrs. P. Jayasree, M.Tech (Assistant Professor, CSE Department) would be very glad to admit the professional suggestions and her unending support to this project. This is done in honor of Dr. Ch. Pavan Sathish, Ph.D (Head of Department, CSE), and Dr. K. Suresh, Ph.D (Principal), Chaitanya Engineering College, Visakhapatnam, which have given the necessary infrastructure and resources. The authors also acknowledge the contribution of the contributors of the Department of Computer Science and Engineering in the academic work.

## REFERENCES

- [1] A. Ahmed, R. Sultana, M. T. R. Ullas, M. Begom, M. M. I. Rahi and M. A. Alam, One machine learning to identify depression and anxiety on the basis of supervised learning, in Proc. IEEE Asia-Pacific Conf. Computer Science and Data Engineering (CSDE), Dec. 2020, doi: 10.1109/CSDE50874.2020.9411642.
- [2] S. C. Guntuku et al., Tracking mental health and symptom mentions on Twitter during COVID-19, *Journal of General Internal Medicine*, vol. 35, no. 9, pp. 2798-2800, Sep. 2020, p. doi: 10.1007/s11606-020-05988-8.
- [3] A. E. Aiello, A. Renson, and P. N. Zivich, Social media- and internet based disease surveillance to support public health, *Annual Review of Public Health*, vol. 41, 101-118, Apr. 2020, doi: 10.1146/annurev-publhealth-040119-094402.
- [4] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, BERT: Pre-training of deep bidirectional transformers to understand language, in Proc. NAACL-HLT 2019, Minneapolis, MN, 2019, pp. 4171-4186.
- [5] R. J. Anandhi, P. Sharma and S. K. S., Chatbot to check the mental health, *International Research Journal of Engineering and Technology (IRJET)*, 2020. [Online]. Available: [www.irjet.net](http://www.irjet.net)The paper has undertaken systematic review of literature on HCI in mental health machine learning to facilitate development of the successful and deployable machine learning systems.
- [6] A. Thieme, D. Belgrave, and G. Doherty, "Machine learning in mental health: Systematic review of the HCI literature to promote the creation of viable and implementable ML systems," *ACM Transactions on Computer-Human Interaction*, vol. 27, no. 5, Article 34, Aug. 2020, doi: 10.1145/33.
- [7] Machine learning techniques to predict physical health diagnoses: A systematic literature review. U. Madububambachu, A. Ukpebor, and U. Ihezue. *Clinical Practice & Epidemiology in Mental Health*, Jul. 2024, doi: 10.2174/011745017931568240607052117.
- [8] Frontiers Review Team, "A review on the machine learning methods of identifying mental health conditions," *Frontiers in Psychiatry*, 2023. [Online]. Available: <https://www.frontiersin.org/articles/mental-health-ml-review>