

Depression And Anxiety Detection Using NLP

Aamina Khan¹ and Dr. Jayasree Ravi²

Department of Computer Science^{1,2}

SVKM's Mithibai College of Arts, Chauhan Institute of Science & Amrutben Jivanlal College of Commerce And Economics (Autonomous), Ville Parle(W), Mumbai, India
khanaamina965@gmail.com

Abstract: *In this study, I have focused on detecting depression and anxiety-related expressions from social media text using Natural Language Processing (NLP). For this purpose, a dataset of Facebook comments collected from mental health support pages was used, as these comments reflect real emotions and experiences shared by users.*

Earlier approaches mostly focused on binary classification. However, in this work, I have extended it to a four-class classification problem, which includes depression, anxiety, both, and neither categories. This helps in understanding In this study, I have considered the problem of detecting depression and anxiety-related expressions based on social media texts with the help of Natural Language Processing (NLP). A dataset of Facebook comments posted on mental health support pages has been employed for this purpose, as the comments present true-to-life text that users share their emotions into.

Most of the previous approaches used binary classification. However, in this work, I have made it a four-class classification problem i.e., with depression, anxiety, both and none categories. This allows for a deeper and more advanced understanding of the expressions of a person.

Different machine learning models like Logistic Regression, Naive Bayes, and Support Vector Machine were employed alongside TF-IDF features. Besides these, a transformer-based model, BERT was also used to grasp better context from the text.

The models were evaluated based on accuracy, precision, recall, and F1-score. The results indicate that BERT outperforms traditional models due to its ability to comprehend the context of multiple sentences. user expressions in a more detailed way..

Keywords: NLP, Depression Detection, Anxiety Detection, Machine Learning, BERT

I. INTRODUCTION

Mental health disorders such as depression and anxiety have become global concerns. World Health Organization (WHO) reports that over 280 million people suffer from depression worldwide, while anxiety disorders affect more than 300 million people. These not only impair the emotional and physical well-being but also academic performance, workplace performance, and interpersonal relationships. Despite the magnitude of these disorders, many people remain undiagnosed, mainly because of the associated stigma, lack of awareness, and inaccessibility of mental health facilities. Reddit, Twitter, Facebook, and other social media platforms have become forums for people to vent their emotions in the digital era. The texts for the study are comments written under posts on Facebook pages aimed at helping people suffering from depression and anxiety. These platforms present a unique opportunity to predict first symptoms of psychological disorders using linguistic data. This can include lexical or sentiment analysis, or an analysis of posting behaviour.

Natural Language Processing (NLP), a subfield of AI, allows machines to interpret and analyze human languages. In this way, the NLP approach to social media data allows revealing various patterns related to mental health issues. The study will evaluate the performance of conventional machine learning models and transformer-based deep learning models for depression and anxiety detection using social media text. In this context, the study is to determine the best



performance between traditional and transformer-based deep learning models in identifying depression and anxiety using social media text data to create automated and scalable solutions that facilitate early intervention in mental health. On top of that, this study also shifts the problem from a simple binary classification to a multi-class problem, which more clearly identifies the type of mental health expression.

II. LITERATURE REVIEW

Author(s) / Year	Model(s)	Key Statement	Research Gap
De Choudhury et al. (2013)	Logistic Regression	Proposed early detection of depression using linguistic and behavioral features from social media achieving ~70% F1-score.	Limited to binary classification and lacks deep contextual understanding.
Yazdavar et al. (2017)	LDA + Sentiment Analysis	Identified depressive symptoms using topic modeling and sentiment-based features.	Does not capture semantic context and relies on traditional techniques.
Orabi et al. (2018)	CNN, LSTM	Applied deep learning models to capture sequential patterns, achieving improved performance over traditional models	High computational complexity and limited interpretability.
Matero et al. (2019)	BERT	Introduced transformer-based model achieving ~86% F1-score with better contextual understanding.	Requires large datasets and high computational resources.
Kumar & Bhattacharyya (2020)	Fine-tuned BERT	Achieved 90–92% F1-score using fine-tuned BERT for mental health detection.	Model complexity and lack of lightweight alternatives.
Patel & Kumar (2021)	CNN-LSTM Hybrid	Combined CNN and LSTM for improved anxiety detection accuracy.	Model is computationally expensive and not scalable.
Zhang & Chen (2022)	Emotion-Aware Transformer	Improved classification by incorporating emotional context into transformer models.	Requires large annotated datasets and lacks real-world deployment.
Singh & Sharma (2023)	Sentiment-Aware BERT	Handled multilingual and code-mixed data effectively using BERT-based model.	Limited generalization across different languages.
Kim et al. (2024)	mhGPT	Proposed lightweight transformer model for efficient mental health detection.	Limited evaluation on diverse datasets.
Kerasiotis et al. (2024)	DistilBERT + Metadata	Combined textual and metadata features to improve detection accuracy (~84.15% F1-score).	Increased complexity and dependency on metadata availability.
Shwetha & Pushpalatha (2024)	Deep Learning Models	Highlighted dominance of transformer-based models in NLP tasks.	Lacks experimental validation.
Teferra et al. (2024)	Ethical AI	Emphasized fairness, transparency,	Limited practical

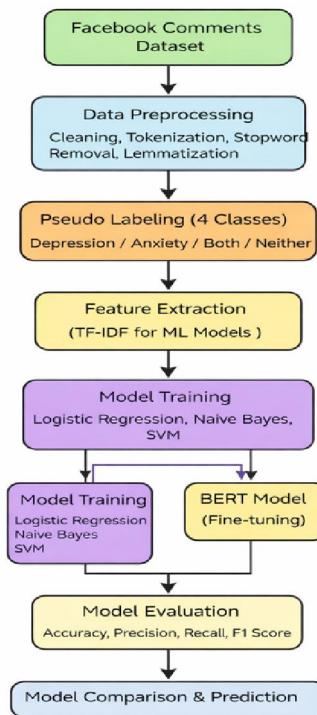


		and cultural sensitivity in AI systems.	implementation.
Villarreal-Zegarra et al. (2024)	NLP Interventions	Showed effectiveness of NLP-based approaches in reducing symptoms.	Does not focus on detection models specifically.
Rohan-Chandrashekar (2024)	NLP-BERT Pipeline	Provided practical implementation of BERT-based NLP system.	Lacks academic evaluation and benchmarking.

III. METHODOLOGY

The system they developed in this study is used to find depression and anxiety expressions from media text using Natural Language Processing techniques. Social media platforms have a lot of things that people write about their feelings and what they go through. This makes social media a good place to look at for health things. The system starts by getting data from pages that help people with health problems. They pick out comments that talk about health. When they have all the data they make it clean by taking out characters, website addresses and common words that do not mean much.

They call this making the data ready to use. Then they put labels on the data so it goes into groups like depression, anxiety, both or none. This helps get the data ready for the computer to learn from. Next they use ways like TF-IDF to change the words into numbers so the computer can understand it. The system uses computer programs, including old ways and new ways like BERT to find patterns in the data.



. Fig 1 represents the overall system workflow

These computer programs are like tools that help find depression and anxiety expressions from media text using Natural Language Processing techniques. Finally, they teach the computer programs. They test them to see how well they work by using things like how accurate they're how precise they are how well they remember and a special score



called F1-score. This step by step way helps the system understand what people are saying accurately and helps find health problems early.

The system developed in this study focuses on finding depression and anxiety-related expressions from media text using Natural Language Processing techniques. It is about using Natural Language Processing techniques to find depression and anxiety expressions, from media text. Fig 1 represents the overall system workflow. It shows all steps from data collection to final model prediction.

A. Data Collection

For this particular study, a dataset of comments from Facebook was used. The dataset consists of comments shared by individuals on mental health support pages. This dataset includes real-world emotions, which makes it suitable for conducting a study on depression and anxiety.

B. Data Preprocessing

Before applying the models, preprocessing of the data was done. The preprocessing of the data includes:

- Removal of URLs, punctuation, emojis, and special characters
- Conversion of text data to lowercase
- Tokenizing sentences into words
- Removal of stopwords
- Lemmatization of words to their base form

These preprocessing steps are necessary to reduce noise in the data and improve the quality of data. In order to understand the data better, the most frequent words in the data were analyzed. Fig 2 is a representation of the top 20 most frequent words found in the data. As observed, words such as “feel”, “time”, and “anxiety” are frequently used, which means that there are emotional content in the data.

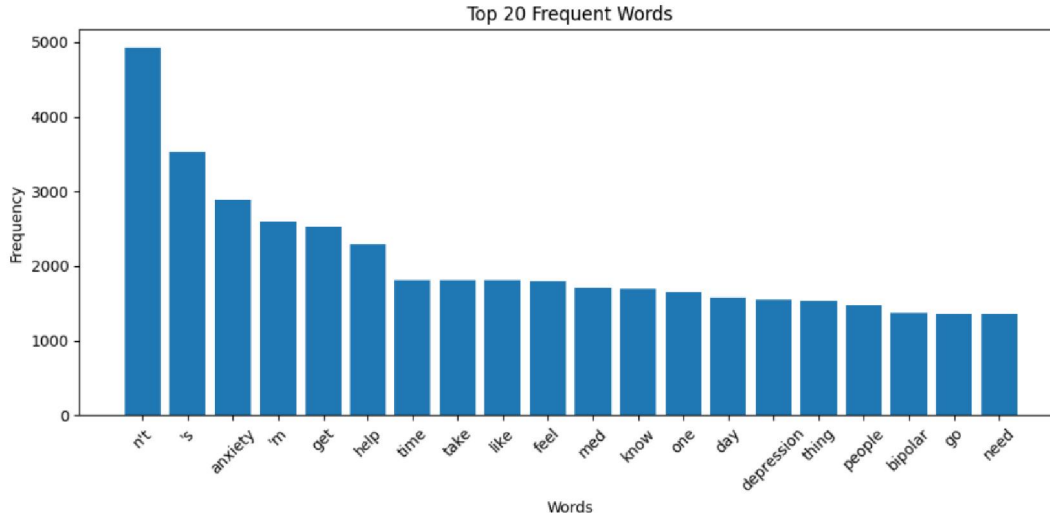


Fig 2 shows the top 20 frequent word

A word cloud was generated to identify the most important words in the data. Fig 3 is a representation of a word cloud created from the data. As observed, larger words represent the frequency of use, showing emotional words used in the data.



E. Machine Learning Models

The following traditional machine learning models were used:

- Logistic Regression
- Naive Bayes
- Linear SVM

These models were trained using the TF-IDF features.

F. BERT Model

In addition to the traditional machine learning models, a transformer-based model was used. The pre-trained BERT model was used for fine-tuning on the dataset for the purpose of multi-class classification. The BERT model is capable of understanding the context of the words used in a sentence. Therefore, the BERT model is more useful for analyzing multiple sentences compared to the traditional TF-IDF-based models.

G. Model Training and Evaluation

The dataset was used for training the models and then evaluating the performance of the trained models. The performance of the trained models was evaluated using the following metrics:

- Accuracy
- Precision
- Recall
- F1 Score

These metrics are used for comparing the various models and finding the best-performing model.

H. Model Comparison

Finally, the performance of the various models was compared with each other. Special emphasis was given to the comparison of the traditional machine learning models with the BERT model

The performance of all models was compared. Figure 5 depicts the comparison of all models based on their performance using F1 score. BERT has been performing best, followed by SVM and then Logistic Regression

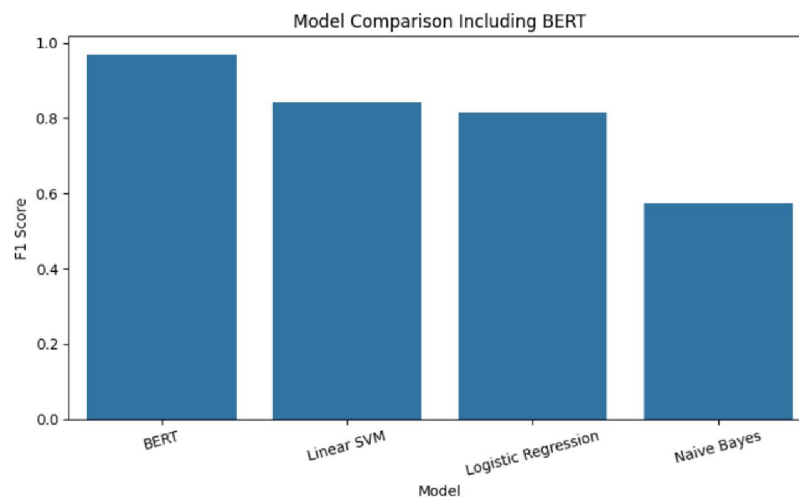


Fig 5 comparison of all models



IV. RESULT

In this study, different models were used to classify social media text into four labels of depression, anxiety, both and neither. Their performance results were measured based on accuracy, precision, recall, and F1 score.

Out of all the models, BERT gave the best result. This figure was much higher compared to traditional models. This is because BERT understands the meaning and context of sentences when the input consists of more than one sentence.

Traditional machine learning models also yielded good results. Linear SVM was the best with an F1 score of about 84.2% followed by Logistic Regression at about 81.4%. Naive Bayes gave the lowest accuracy with an F1 score of around 57.3%. TF-IDF is used by these models, which are limited to word frequency-based understanding and cannot make meaningful observations based on the text.

From these results, we can say that models which understand context of the text perform better on this type of problem. BERT helps in working better because the context of sentences can be understood and the relationship of words and sentences can be understood in a better way rather than words being treated individually and based on pattern worked on, which is the case for traditional models.

V. CONCLUSION

This study shows in conclusion that the use of NLP techniques to detect mental health-related expressions can be done using social media textual evidence. Coupled with mental health-related expressions, extending the problem to four-classes helped in giving much meaning and useful results as opposed to simple binary classification.

However, it is important to note that the labels in this study were created through a rule-based method, rule says that Rule-based labelling was used in which keywords related to depression and anxiety were already decided and compared with the text. Thus, each comment was labeled as depressed, anxious, both, or neither. That means the model only sees patterns in keywords and not real medical conditions. Hence, the results should not be considered as medical diagnosis.

Traditional models are useful as well as deep models, however, BERT deep model performs better in understanding the context of text related to mental health, overall, the results obtained from this study show that.

ACKNOWLEDGMENT

I would like to express my sincere gratitude to my project guide Dr. Jayshree Ravi for their support and guidance throughout this research work.

REFERENCES

- [1] M. De Choudhury et al., "Predicting depression via social media," in Proc. AAAI Int. Conf. Web Social Media, 2013.
- [2] A. Yazdavar et al., "Analysis of depressive symptoms from Reddit data," in Proc. IEEE ASONAM, 2017.
- [3] H. Orabi et al., "Deep learning approaches for depression detection," in Proc. ACL, 2018.
- [4] M. Matero et al., "Risk of depression detection using BERT," in Proc. ACL, 2019.
- [5] Patel et al., "CNN-LSTM hybrid model for anxiety detection," *Procedia Computer Science*, 2021.
- [6] Zhang et al., "Emotion-aware transformer for mental health detection," *IEEE Xplore*, 2022.
- [7] Singh et al., "Sentiment-aware BERT for code-mixed data," *ScienceDirect*, 2023.
- [8] Teferra et al., "Ethical and explainable AI in mental health," *i-JMR*, 2024.
- [9] Villarreal-Zegarra et al., "NLP-based interventions for mental health: A meta-analysis," *JMIR*, 2024.
- [10] Shwetha et al., "Deep learning trends in NLP," Springer, 2024.
- [11] Kumar et al., "From machine learning to GPT: Evolution of NLP," *CFILT, IIT Bombay*, 2024.
- [12] R. Chandrashekar, "NLP-BERT pipeline implementation," *GitHub Repository*, 2024.
- [13] Kim et al., "mhGPT: Mental health detection using large language models," *arXiv*, 2024.
- [14] Kerasiotis et al., "DistilBERT with metadata for mental health detection," Springer, 2024.

