

AI-Based Mental Health Emotion Detection from Text using NLP

Submitted by: **Vaishnavi Pawar, Mahendra kondekar**
Marathwada Institute of Technology, Chhatrapati Sambhajnagar

Abstract

Mental health issues such as anxiety, depression, and suicidal tendencies are rapidly increasing in today's digital era. With the growth of social media and online communication platforms, individuals frequently express their thoughts and emotions through text-based interactions. These textual expressions contain valuable psychological indicators that can be analyzed using Artificial Intelligence. This research presents an AI-based system that uses Natural Language Processing (NLP) and Machine Learning techniques to detect emotional states from user-generated text. The system classifies text into four major categories: Normal, Anxiety, Depression, and Suicidal. A dataset of approximately 9,916 labeled text samples was used for training and evaluation. Various preprocessing techniques such as tokenization, stopword removal, and lemmatization were applied to clean and standardize the data. TF-IDF vectorization was used for feature extraction, and Logistic Regression provided the best accuracy of around 85%. The system is deployed using a Streamlit web application that allows real-time prediction and suggestions. This research demonstrates how AI can support early mental health detection and awareness, providing a scalable and accessible solution for emotional well-being monitoring.

Introduction

Mental health has become one of the most critical issues in modern society. With increasing stress, competition, and digital exposure, many individuals experience emotional distress but hesitate to seek professional help. Traditional mental health assessment methods, such as clinical interviews and psychological tests, are often expensive, time-consuming, and not easily accessible to everyone. At the same time, people are actively expressing their emotions through social media platforms, blogs, and messaging applications. These digital expressions provide an opportunity to analyze emotional states using data-driven approaches. Natural Language Processing enables machines to understand human language and extract meaningful patterns from text. By combining NLP with Machine Learning, it is possible to develop systems that can automatically detect emotional conditions. This research focuses on building such a system that can analyze user-written text and classify it into emotional categories. The goal is to create an accessible tool that helps in early identification of mental health issues and promotes awareness.

Objectives

The primary objective of this research is to design and develop an intelligent system capable of detecting mental health conditions from textual data. The system aims to classify user input into four emotional categories: Normal, Anxiety, Depression, and Suicidal. Another important objective is to apply Natural Language Processing techniques to clean and process textual data efficiently. The research also aims to implement machine learning algorithms and evaluate their performance to achieve high accuracy. Additionally, the project focuses on developing a user-friendly web application that provides real-time predictions and suggestions. The ultimate goal is to contribute to mental health awareness by providing a scalable, accessible, and AI-driven solution.

Study Area

This research lies at the intersection of Artificial Intelligence, Natural Language Processing, and Mental Health Analysis. The study focuses on analyzing text data collected from online platforms such as social media,

forums, and public datasets. These sources provide a rich collection of human expressions that reflect emotional states. The study explores how linguistic patterns, word usage, and sentence structures can indicate psychological conditions. By applying machine learning techniques, the system learns to recognize these patterns and classify them into predefined categories. This study contributes to computational mental health research and demonstrates how AI can be used for emotional analysis in real-world scenarios.

Tools and Technologies

The project is implemented using Python, which is widely used for data science and machine learning applications. Natural Language Processing libraries such as NLTK and spaCy are used for text preprocessing tasks including tokenization, stopwords removal, and lemmatization. Scikit-learn is used for implementing machine learning algorithms such as Logistic Regression, Naïve Bayes, and Support Vector Machines. Data visualization tools like Matplotlib and Seaborn are used to analyze results and generate graphs. Streamlit is used to develop an interactive web application that allows users to input text and receive predictions in real time. The development environment includes Jupyter Notebook for experimentation and Visual Studio Code for application development.

Dataset

The dataset used in this research consists of approximately 9,916 text samples collected from various online sources. Each sample is labeled into one of four categories: Normal, Anxiety, Depression, and Suicidal. The dataset includes a mix of short and long text entries to simulate real-world scenarios. Data preprocessing plays a crucial role in ensuring the quality of the dataset. Techniques such as removing special characters, eliminating stopwords, and applying lemmatization are used to clean the data. The dataset is then transformed into numerical form using TF-IDF vectorization, which helps the machine learning model understand the importance of words in the text.

Methodology

The methodology of this research involves several steps, starting with data preprocessing. Raw text data is cleaned and standardized using NLP techniques. The cleaned text is then converted into numerical features using TF-IDF vectorization. The dataset is divided into training and testing sets to evaluate model performance.

Multiple machine learning algorithms are implemented and compared. Logistic Regression is selected as the best-performing model.

Results and Discussion

The results of the research show that the proposed system achieves an accuracy of approximately 85%. The model performs well in distinguishing between Normal and Suicidal categories, while some confusion is observed between Anxiety and Depression due to similar linguistic patterns. The confusion matrix and evaluation metrics such as precision, recall, and F1-score are used to analyze performance.

The integration of the model into a web application enhances usability and demonstrates real-time capabilities. Overall, the results indicate that NLP and machine learning can effectively detect emotional states from text and provide valuable insights into mental health conditions.

Conclusion

This research successfully demonstrates the application of Artificial Intelligence and Natural Language Processing in detecting mental health conditions from text. The system provides a fast, accessible, and scalable solution for analyzing emotional states. While it does not replace professional diagnosis, it serves as an effective preliminary screening tool. The integration of machine learning with a user-friendly interface makes the system practical for real-world use. The study highlights the potential of AI in improving mental health awareness and early intervention. and scalable solution for analyzing emotional states. While it does not replace professional diagnosis, it serves as an effective preliminary screening tool.

Future Work

Future enhancements of this project may include the use of advanced deep learning models such as BERT and transformer-based architectures to improve accuracy. The system can be extended to support multiple languages, making it accessible to a wider audience. Integration with chatbot systems can provide real-time conversational support. Additionally, developing a mobile application can increase usability and accessibility. Further research can explore multimodal analysis by combining text, voice, and facial expressions for a more comprehensive understanding of emotions.

References

1. Jurafsky, D., & Martin, J. H. (2023). *Speech and Language Processing*.
2. Bird, S., Klein, E., & Loper, E. (2009). *Natural Language Processing with Python*. O'Reilly Media.
3. Pedregosa, F. et al. (2011). *Scikit-learn: Machine Learning in Python*. *Journal of Machine Learning Research*.
4. Devlin, J. et al. (2019). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding.
5. Streamlit Official Documentation.
6. Kaggle Mental Health Text Classification Datasets.