

## Exploring NLP for Mental Health Insights: Multi-Class Classification of Online Forum Texts

**Authors :** Jennifer Patricia

**Abstract :** With the increasing incidence of mental health issues, there is a real need for early detection, which is currently limited by stigma and ignorance. This study attempts to explore multi-class classification models to analyze mental health problems through social media texts. The goal of the classification model is to categorize text into one of six categories of mental health problems and thus to provide patterns of the language which might serve as an early indication of these problems. After data collection and labeling, the dataset was resampled to balance the dataset for model training. Some of the important steps for data preprocessing included tokenization, the removal of unnecessary characters and labels, and one-hot encoding. To further understand the language used in expressing the different conditions, word clouds and bigram analyses were conducted. The models used for the first training are BERT + XGBoost, T5, and MentalBert. The final results demonstrated that T5 and MentalBERT achieved the highest accuracy of 0.83, significantly outperforming BERT + XGBoost, which obtained an accuracy of 0.6.

**Keywords :** mental health detection, exploratory data analysis, natural language processing, multi-class classification, data preprocessing, BERT, XGBoost, T5, MentalBERT

**Conference Title :** ICCST 2025 : International Conference on Computer Science and Technology

**Conference Location :** Bali, Indonesia

**Conference Dates :** July 10-11, 2025