

# Final Project Report:

## Mental Health Prediction

### 1. Abstract and Introduction:

This paper investigates the use of a pre-trained BERT (Bidirectional Encoder Representations from Transformers) model to classify text into health conditions. With the growing demand for mental health awareness, especially in identifying distress through text, this project aims to leverage natural language processing (NLP) for predicting health-related states from textual data. The model uses multi-label classification to predict the health condition of a given piece of text. By employing a weighted BERT model, the approach handles class imbalance in the data effectively. The results demonstrate the model's efficiency in predicting multiple health conditions with high accuracy.

Mental health is an increasingly significant global concern, and understanding people's health-related issues through text has become an innovative approach for healthcare systems. This project focuses on building an automatic classification system that predicts specific health conditions from written text. Using the BERT model, the model is fine-tuned to handle the task of predicting health conditions such as PTSD, anxiety, homelessness, and stress. This approach can be applied in various mental health applications, such as chatbots or automated systems to detect potential health risks based on user input.

### 2. Prior Related Work

Previous studies have demonstrated the effectiveness of deep learning models, particularly BERT, for health-related text classification. The BERT model, with its ability to process language context in both directions, has been extensively used in various NLP tasks. While research into emotion recognition from text is extensive, using NLP models to detect specific health conditions like PTSD, anxiety, and homelessness from text is less explored. Some studies (e.g., Zhang et al., 2020) have shown that deep learning models can perform well on such tasks when trained on large annotated datasets, but there is little work focusing on fine-tuning BERT for a multi-label health condition prediction task.

### 3. Methodology/Model

The approach uses BERT-based architecture for sequence classification. The model is fine-tuned on a labelled dataset containing text samples, each with a label for a health condition. The following steps were undertaken:

1. **Data Preprocessing:** The text data was labelled with various health conditions (e.g., PTSD, anxiety, homelessness). These labels were then mapped to numerical values suitable for training the BERT model.
2. **Model Architecture:** The model was extended from the original BERT architecture to output a set of logits representing the likelihood of each health condition. The final output of the model was the health condition logits, which were later used for loss computation.
3. **Training:** The model was trained using a cross-entropy loss function, with class weights applied to account for any imbalance in the data across health conditions.
4. **Fine-tuning:** The pre-trained BERT model was fine-tuned for three epochs with a batch size of 8, using the Adam optimizer with weight decay. Evaluation was done using accuracy for health condition classification.

### 4. Experiments and Results

#### Dataset:

The dataset consisted of text samples labelled with health conditions (such as PTSD, homelessness, anxiety). The dataset was divided into training and test sets with an 80-20 split to evaluate model performance.

#### Training Configuration:

- **Model:** BERT-base-uncased
- **Epochs:** 3
- **Batch Size:** 8
- **Learning Rate:** Adam optimizer with weight decay and warmup steps
- **Loss Function:** Weighted cross-entropy loss
- **Class Weights:** Health classes were assigned weights to handle class imbalance.

## Results:

The model achieved an accuracy of 80% in predicting health conditions, indicating the effectiveness of the BERT model in handling the health classification task.

Metric	Health Accuracy
Accuracy	80%

The use of class weights helped to improve the model's performance, particularly for minority classes.

## 5. Analysis & Conclusion

This project demonstrates the potential of using BERT for health condition classification. By leveraging BERT's pre-trained language understanding and fine-tuning it on a specific health-related dataset, the model effectively predicts various health conditions from text. The accuracy achieved shows that the model is capable of making reliable predictions, making it a useful tool for automated mental health detection and support.

Future improvements can include expanding the dataset, exploring more diverse health conditions, and optimizing hyperparameters. Additionally, further research into multi-task learning could help in improving the overall performance of models tasked with multiple health-related predictions.

## 6. Demo/System (if any)

At this stage, no live demo or system is available. However, the model is designed to be integrated into applications such as chatbots or automated systems where users provide text, and the model predicts their health condition. The model can be further developed into a user-friendly interface for mental health applications.