MENTAL HEALTH PREDICTION USING TRANSCRIPT DATA

Praveen. T
23MDT0051

ABSTRACT

This research aims to create an intelligent system for mental health prediction based on treatment session transcript data, utilizing Natural Language Processing (NLP) and deep learning techniques. The process begins with data preparation, which involves cleaning, tokenizing, and preparing raw transcript text to meet the input requirements of transformer-based language models. To capture the complex language patterns associated with diverse mental health problems such as depression, anxiety, or PTSD, we use a pre-trained model called BERT (Bidirectional Encoder Representations from Transformers). This model outperforms previous models in terms of comprehending linguistic context and semantic structure.

Following initial deployment, the model is assessed on a dataset to determine its baseline performance. It is then fine-tuned using particular mental health transcript data to improve accuracy and domain flexibility. The AdamW optimizer improves the fine-tuning process by detaching weight decay from gradient updating, resulting in improved generalization and faster convergence.

Following optimization, the model is used to forecast mental health states based on fresh user inputs. This prediction phase delivers insights based on the emotional and psychological patterns shown by the transcript text. The suggested system provides a reliable and scalable technique to early diagnosis of mental health disorders, with applications in clinical support systems, telemedicine platforms, and treatment monitoring tools. It bridges the gap between unstructured linguistic data and actionable mental health insights, hence promoting proactive mental wellness treatments.

Keywords: BERT (Bidirectional Encoder Representations Form Transforms, PTSD (Post Traumatic Stress Disorder), AdamW (Adaptive Moment Estimation with Decay)

Chapter 1

Introduction

1.1 OBJECTIVES

The objective of this project is to create a model that is used to predict mental health using transcript data. This model is processed with a BERT pre-trained model for text classification to find mental health conditions like depression, anxiety, etc. With the help of the AdamW optimizer, is used to improve the model's performance during the training process. By analyzing the emotions and patterns in the transcript, the project aims to predict mental health and assist in understanding mental health conditions more easily and effectively.

1.2 BACKGROUND

Nowadays, almost everyone has some mental health disorders, like depression, anxiety, and post-traumatic stress disorder (PTSD), which have become a world concern that affects millions of people. Traditional diagnostic techniques frequently rely on the subjective evaluation of mental health professionals. With advancements in Natural Language Processing, automated systems can analyze text-based conversations, such as therapy transcripts, to more accurately identify mental health issues.

1.3 MOTIVATION

Mental health issues are frequently undiagnosed due to a lack of access to professionals, long wait times, social stigma, and high treatment costs. As text-based therapy and online support platforms gain popularity, there is a high demand for automated tools that can aid in early detection. Early intervention leads to better treatment outcomes and overall emotional well-being. Natural language processing (NLP) advances now allow for the analysis of language patterns in therapy session transcripts, which can help identify signs of mental health issues. Automated systems can reduce clinician workload, increase accessibility, and promote timely, personalized mental health care for a wide range of populations.

1.4 PROJECT STATEMENT

The aim of this project is to develop a model to predict mental health using therapy transcript data. The BERT (Bidirectional Encoder Representations from Transformers) pre-trained model was used to fine-tune the classification of mental health conditions, and the model was optimized using the AdamW (Adaptive Moment Estimation Weight Decay) optimizer for better performance by analyzing the text and predicting the mental health condition.

1.5 EXISTING SYSTEM

Natural language processing (NLP) is a valuable tool for analyzing trends within therapy transcripts. The existing system uses NLP techniques to gain insights from textual data and predict mental health conditions and patterns. That approach involves analyzing therapy transcripts and classifying the patterns as Early signs of Depression and severe after-effects of Prolonged Depression, as per the patient responses.

In this system, they use traditional machine learning classifier such as Naive Bayes, Support Vector Machine, and logistic regression. For converting the text into numerical features, they used vectorization techniques like TF-IDF and Count Vectorizer. These models help in identifying depression-related patterns by evaluating the linguistic expressions used in the therapy sessions.

1.6 SCOPE OF THE PROJECT

The scope of this research is to create a classification model based on natural language processing (NLP) that uses transcripts of treatment sessions to predict mental health issues. In order to detect emotional and psychological signs in text, this study makes use of a pre-trained BERT model that has been refined using labeled transcript data. By providing an extra tool for the early detection of disorders like anxiety and depression, the system seeks to support mental health providers. The model performs better when the AdamW optimizer is used, which improves its ability to generalize on unknown inputs. The project's goal is to provide experts with data-driven insights, not to replace clinical diagnosis. For real-time monitoring and support, it can be further expanded to other mental health categories and incorporated into digital therapy platforms.

Chapter 2

Literature Survey

2.1 SUMMARY OF THE EXISTING WORKS

Pilot research was done to investigate the annotation process for mental health transcripts, with the goal of assisting in the creation of credible datasets for training machine learning models. The study involves expert clinicians and unskilled participants annotating transcripts in two phases: before and after annotation training. The results revealed low to moderate inter-rater reliability across doctors and subjects, with improvements noted following training. The findings emphasize the need of educating annotators and establishing unambiguous categorization labels in order to improve annotation quality. This study helps to build a more efficient and consistent data gathering approach for AI-powered mental health transcript analysis. [1]

This study focuses on the function of Natural Language Processing (NLP) in recognizing and categorizing mental health disorders from therapy transcripts. The suggested system classifies text into 'Early indications of depression' and 'Serious after-effects of chronic depression' utilizing classifiers such as Naïve Bayes, Support Vector Machine, and Logistic Regression, coupled with TF-IDF and Count vectorization algorithms. By studying language patterns in therapy talks, the technology hopes to help both patients and therapists recognize and comprehend mental health issues earlier. This method not only helps with diagnostic and predictive analysis, but it also improves the quality and efficacy of counseling and therapy sessions. [2]

This systematic review looks at recent advances in using machine learning to predict mental health disorders. The authors followed the PRISMA approach to examine and analyze 30 selected research papers from reliable databases, classifying them by mental health conditions such as schizophrenia, bipolar disorder, anxiety, depression, PTSD, and child mental health. The study examines major results, frequent obstacles, and constraints encountered by researchers in this field. It also makes recommendations for future study, highlighting the need of improved methodology and robust data in improving the efficacy of machine learning applications in mental health prediction and care. [3]

This paper presents a supervised machine learning strategy that employs Support Vector Machines (SVM) to provide a blood-based diagnostic tool for Major Depressive Disorder (MDD). The model attained a cross-validated sensitivity and specificity of 90.6% while assessing transcriptome data from peripheral blood samples of 32 MDD patients and 32 matched controls using a panel of ten transcripts. In addition, a logistic equation was used to calculate the likelihood of depression, allowing for individualized diagnosis by modifying sensitivity and specificity per individual. This study demonstrates the potential for objective, transcript-based testing to enhance MDD diagnosis and individualized psychiatric therapy. [4]

The growing subject of 'predictive analytics in mental health' has lately sparked widespread interest, with the bold promise of revolutionizing clinical practice in psychiatry, akin to breakthroughs in customized and precision medicine. We present an overview of the field's key questions and challenges, with the goal of (1) proposing general guidelines for predictive analytics projects in psychiatry, (2) providing a conceptual introduction to core aspects of predictive modeling technology, and (3) fostering a broad and informed discussion involving all stakeholders, including researchers, clinicians, patients, funding bodies, and policymakers. [5]

This project investigates the creation of an end-to-end clinical transcription tool designed to automate EHR recording for mental health professionals. Using 65 simulated patient-provider transcripts, the researchers refined a transformer-based language model for extracting and categorizing important information. The algorithm performed well in data extraction (F1 = 0.94) but has room for improvement in EHR category categorization (F1 = 0.18). A rule-based module was utilized to create organized clinical notes from the retrieved data. The suggested pipeline has the potential to cut documentation time by 70-80%, reducing clinician stress and increasing the efficiency of mental health treatment delivery. [6]

This paper describes a modular deep learning pipeline for detecting depression severity based on speech transcription data. The study compares multiple deep learning architectures using the E-DAIC dataset, demonstrating that temporal pooling of latent representations—without relying on temporal dynamics—outperforms recurrent neural networks. The suggested strategy improved the Concordance Correlation Coefficient (CCC) by 8.8% over existing state-of-the-art methodologies, illustrating the efficacy of simplified representations in depression evaluation. [7]

This systematic review examines the use of Natural Language Processing (NLP) in assessing Mental Health Interventions (MHI) by examining 102 papers drawn from a pool of 19,756 articles. The review adheres to PRISMA principles and examines computational methodologies, clinical applications, and limits in NLP-MHI research. The results reveal a fast increase in research since 2019, driven by bigger sample sizes and the use of large language models. Textual characteristics were more predictive than aural features. Studies frequently concentrated on patient presentation, intervention response, monitoring, and provider dynamics. Linguistic bias, insufficient repeatability, and underrepresentation are among the most significant difficulties observed. The authors suggest the NLPxMHI research paradigm to bridge these gaps while also improving the therapeutic usefulness and fairness of NLP applications in mental health. [8]

This study uses clinical NLP to solve data imbalance in PTSD diagnosis, presenting two unique text augmentation frameworks based on Large Language Models (LLMs). The authors improve the Extended Distress Analysis Interview Corpus (E-DAIC) using a zero-shot (ZS) strategy, which generates synthetic PTSD transcripts, and a few-shot (FS) approach, which rephrases existing samples. Both strategies increase diagnostic performance, but the ZS method with GPT embeddings has the best accuracy. The enhanced data closely mimics the original corpus, providing a scalable and cost-effective alternative for extending clinical NLP datasets. This study demonstrates the potential of LLMs in producing high-quality synthetic data for mental health applications. [9]

This large-scale study evaluated data from 68,894 individuals from 24 countries to investigate determinants of traumatic event exposure, which is a relatively understudied subject despite its well-documented health implications. Over 70% of respondents said they had experienced at least one traumatic incident, with five kinds accounting for more than half of all exposures. Interpersonal violence was found to be a substantial predictor of future trauma, but marriage was consistently protective. The findings imply that preventative treatments should target persons at higher risk of recurrent trauma, particularly victims of interpersonal violence, in order to attenuate long-term health consequences. [10]

Designed for quick and non-intrusive mental disease screening with both active (voice recordings) and passive (digital phenotypic) methods. The EMU app gathered data and used anxiety and depression screening questions to classify it. The results revealed that participants preferred providing scripted audio versus passive data. Machine learning models trained on audio produced good results, with F1-scores of 0.746 for depression, 0.667 for anxiety, and 0.706 for suicide thoughts utilizing scripted recordings. Jitter and MFCCs were important elements, with unscripted audio showing a high frequency of help-related words. The EMU framework is a scalable, unbiased alternative to standard mental health screening that includes a publicly available dataset for future study. [11]

This research focuses on predicting mental health disorders using machine learning approaches applied to publicly available, label-encoded data. The project's goal is to identify mental health disorders such as stress, anxiety, and depression in adults over the age of 18, and then incorporate the prediction model into a website for real-time user assessments.[12]

Recent advances in informatics, notably artificial intelligence, machine learning, and deep learning, are revolutionizing brain and mental health research. These technologies aid precision psychiatry by allowing for tailored detection, diagnosis, and treatment of mental health problems such as depression. While AI systems show potential throughout the treatment process and enable a shift toward evidence-based, tailored care, existing models lack empirical validation and better patient outcomes. Future research must prioritize clinical validation, multidisciplinary collaboration, and access to different datasets for wider use.[13]

This research examines current machine learning algorithms for detecting ADHD and depression, two increasingly common mental health problems. It emphasizes the importance of many modalities in training machine learning models, including fMRI, EEG, medical notes, video, and voice. Given the demand on mental health facilities and the trend toward remote consultations, the study underscores the rising need for AI-based diagnostic tools to help with clinical decision-making and enhance access to therapy.[14]

This study examines the many AI and machine learning strategies used to diagnose depression and analyze emotions. It investigates using facial expressions, emotional chatbots, and social media writings to detect mood and despair. Naive Bayes, SVM, LSTM-RNN, Logistic Regression, and ANN are used to recognize emotions in both text and pictures. The report underlines the rising demand for automated depression detection systems across age groups, as well as the ongoing research hurdles in this sector.[15]

Chapter 3

Requirements

3.1 HARDWARE REQUIREMENTS

Processor	Intel i5
RAM	At least 16GB
Hard disk	100GB or more
GPU	NVIDIA T4 GPU
NETWORK	Stable internet Connection

Table 3.1 Hardware Requirements

3.2 SOFTWARE REQUIREMNETS

Operating system	Windows10/Windows11
Coding Language	Python 3.10+
Libraries & Frameworks	• Pandas • NumPy • Sklenar • Matplotlib • Seaborn • Troch • Re • Nltk • Datasets • Spacy • ipywidgets

Table 3.2 Software Requirements

3.3 DEVELOPMENT ENVIRONMENT

Code Editor	Google Colab
Alternatives	Jupyter Notebook, and VS code

Table 3.3 Development Environment

Chapter 4
Project Planning

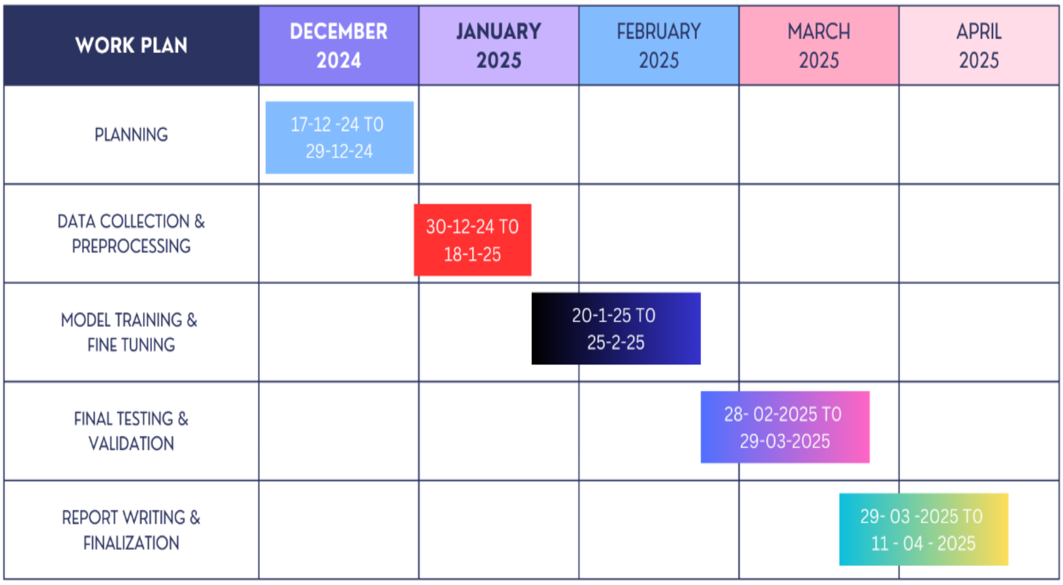
4.1SCHEDULE TIME TABLE

S.NO	MONTH -WEEK	PLAN
1	DECEMBER - WEEK 3 To 4	Identification the problem statement & Gathering the literature reviews.
2	JANUARY- WEEK 1	Finalize project objectives
3	JANUARY- WEEK 2 To 4	Collecting data and Data preprocessing
4	FEBRUARY- WEEK 1 To 4	Define the methodology and build the model based on the prepared outcome
5	MARCH – WEEK 1 To 2	Testing and fine turning model
6	MARCH – WEEK 3 To 4	Final Testing and validation
7	APRIL - WEEK 1 To 2	Draft report and documentation
8	APRIL - WEEK 3	Discussion with guide about result and finalize the report
9	APRIL - WEEK 4	Report review
10	MAY - WEEK 1	Final review and project presentation

Table 4.1 Schedule time table

The above (table 3.4) shows the schedule time table of the project, this includes planning and project proceeds in a disciplined, well-planned way. Also attached related Gantt chart in fig 3.1.

4.2GANTT CHART



WORK PLAN	DECEMBER 2024	JANUARY 2025	FEBRUARY 2025	MARCH 2025	APRIL 2025
PLANNING	17-12 -24 TO 29-12-24				
DATA COLLECTION & PREPROCESSING		30-12-24 TO 18-1-25			
MODEL TRAINING & FINE TUNING			20-1-25 TO 25-2-25		
FINAL TESTING & VALIDATION				28- 02-2025 TO 29-03-2025	
REPORT WRITING & FINALIZATION					29- 03 -2025 TO 11 - 04 - 2025

fig 4.1 Gantt Chart

The above (fig 3.1) shows the Gantt chart of the system, this includesplanning, Data Collection and preprocessing, Model Training and fine turning, implementation and testing, and reporting writing.

Chapter 5

Analysis & Design

5.1PROPOSED METHODOLOGY

The Proposed system aims to find mental health conditions using transcript data by applying advanced Natural Language Processing techniques. Instead of using traditional models that mostly rely on basic feature extraction and machine learning algorithms, this system used a pre-trained BERT model. BERT can understand the meaning of words within a sentence, making it well-suited for analyzing more complex emotional and psychological language.

The treatment transcripts are first cleaned, tokenized, and formatted to the BERT-specific structure. The model is then fine-tuned with labeled data to categorize the text into mental health-related categories such as anxiety and depression. The AdamW optimizer is used to enhance performance and ensure efficient learning by improving model convergence and preventing overfitting.

The system is tested using conventional classification measures such as accuracy, precision, recall, and F1-score to ensure its reliability and efficacy. This technique outperforms classic classifiers like Naïve Bayes or SVM in terms of accuracy, context, and scalability. The suggested method is intended to help mental health practitioners by providing an extra layer of analysis that can aid in early diagnosis and intervention without replacing clinical judgment.

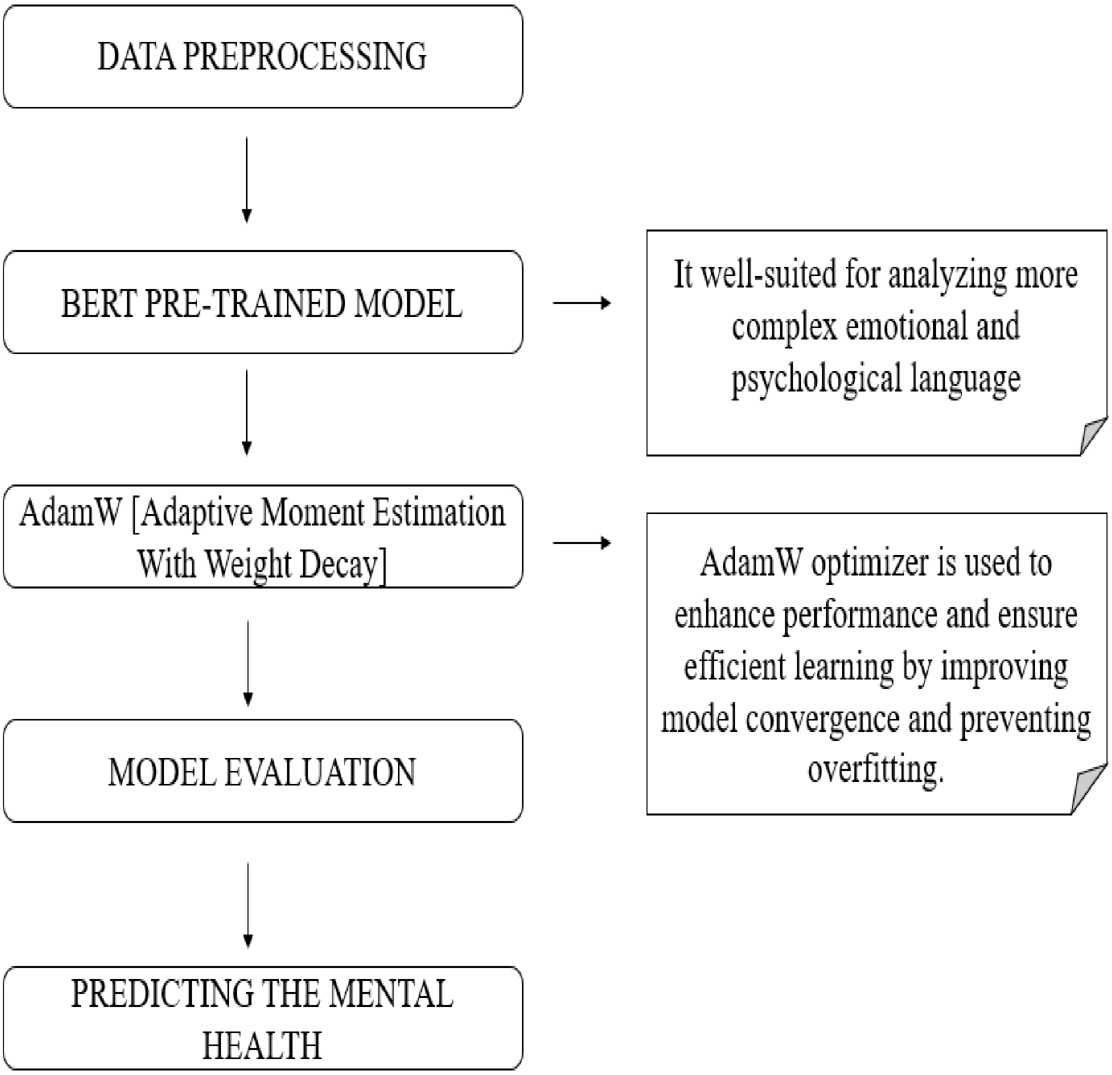

DATA PREPROCESSING
BERT PRE-TRAINED MODEL
It well-suited for analyzing more complex emotional and psychological language
AdamW [Adaptive Moment Estimation With Weight Decay]
AdamW optimizer is used to enhance performance and ensure efficient learning by improving model convergence and preventing overfitting.
MODEL EVALUATION
PREDICTING THE MENTAL HEALTH

fig 5.1 block diagram about how the system flows

5.2 ADVANTAGES OF PROPOSED METHODOLOGY

BERT captures the contextual meaning of words by taking into account both left and right word surrounds, making it extremely useful for comprehending emotional nuance and psychological intricacies in therapy transcripts.

Pre-trained transformers, such as BERT, outperform classical models (e.g., Naïve Bayes, SVM) by utilizing deep semantic links, resulting in greater accuracy, precision, recall, and F1 scores.

Unlike traditional machine learning models that need manual feature selection and extraction (e.g., TF-IDF, n-grams), BERT operates on raw text with minimum preparation, simplifying the pipeline and improving performance.

BERT's pre-training on large datasets makes it applicable to a wide range of disciplines. Fine-tuning enables it to understand domain-specific variations in mental health without requiring large labeled datasets.

Therapy transcripts frequently feature nonlinear, emotionally charged talks. BERT can grasp such complexity better than bag-of-words or rule-based models.

The AdamW optimizer improves the model's generalization, reduces overfitting, and speeds up convergence, all of which are critical for real-world mental health applications.

Fine-tuning enables the model to classify numerous mental health problems (e.g., anxiety, depression, and PTSD), which would be difficult for simpler classifiers.

5.3 SYSTEM ARCHITECTURE

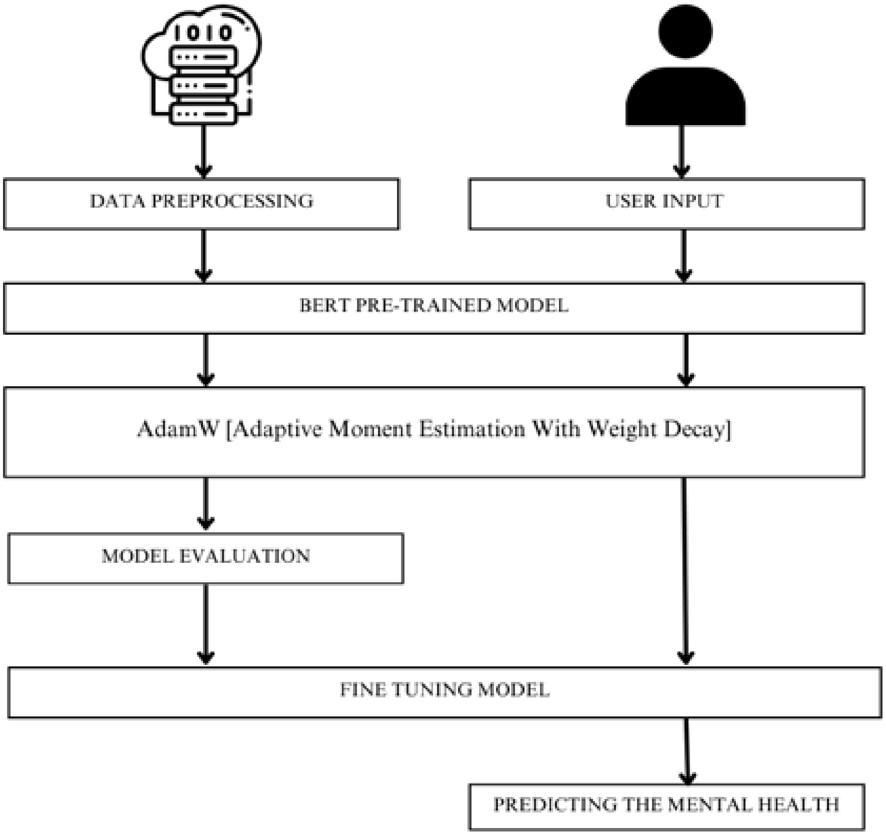


fig 5.2 architecture diagram

The system architecture diagram depicts a workflow for predicting mental health with a BERT-based NLP model. The process begins with data preparation, which involves cleaning and preparing raw text data. This processed data is then supplied into a pre-trained BERT model that recognizes the context of the text. The model is trained and optimized with the AdamW optimizer, which improves learning efficiency and reduces overfitting. Following evaluation, the model is fine-tuned for increased accuracy. On the other hand, user input is analysed using the same BERT model to create predictions. The end result is a categorization of the user's mental health status based on the text input.

5.4 UML DIAGRAMS

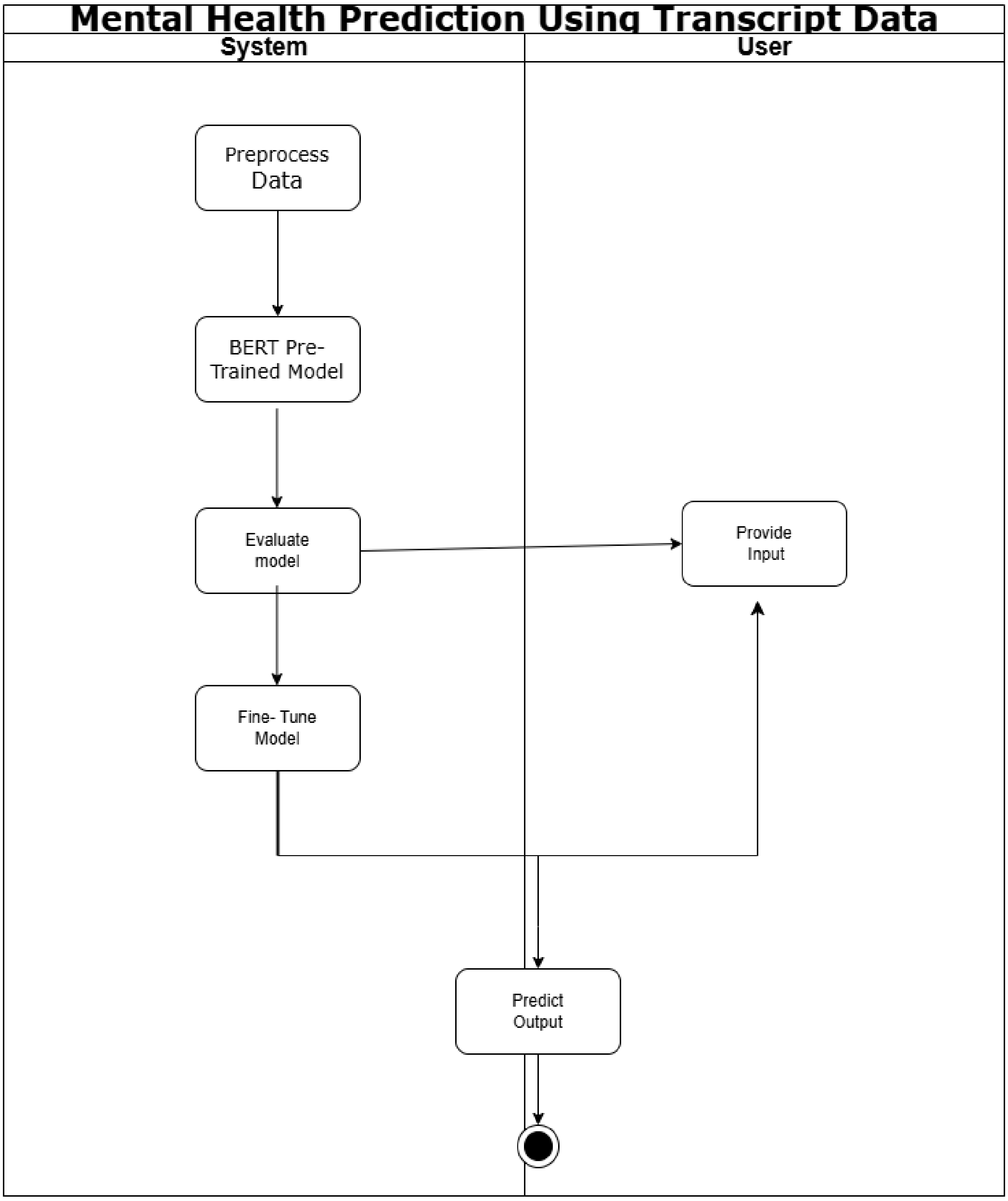


fig 5.3 activity diagram for model

The activity diagram (fig 4.3) named "Mental Health Prediction Using Transcript Data" depicts the system's workflow and user interactions. The system starts by preprocessing the transcript data, then uses a BERT pre-trained model for assessment. Based on the evaluation, the model is fine-tuned to increase performance. Meanwhile, the user supplies input data (for example, fresh transcript text), which the algorithm uses to forecast mental health outcomes. This organized approach guarantees that data is handled efficiently, models are optimized, and mental health disorders are accurately predicted based on transcripts supplied by users.

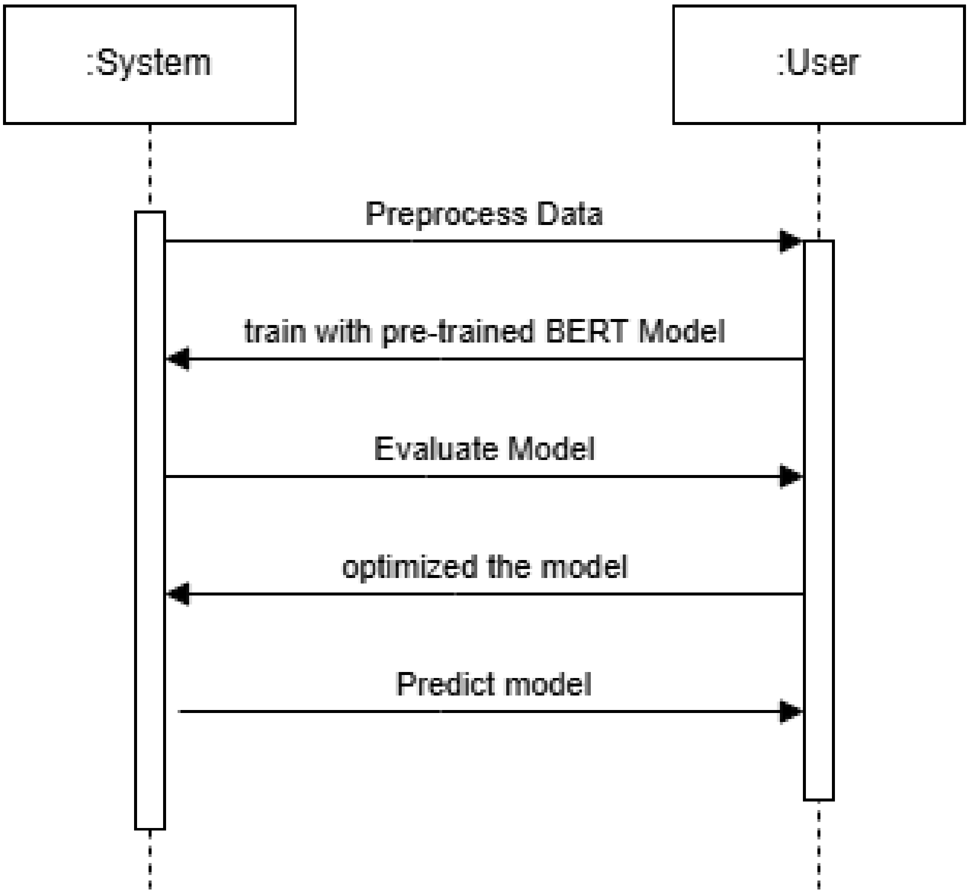


fig 5.4 sequence diagram for model

The sequence diagram (fig 4.4) depicts the interaction of the user and the system in the mental health prediction process. It starts with data preparation, then applies a pre-trained BERT model to the input data. The model is then assessed and improved using the AdamW optimizer to improve performance. Finally, the algorithm estimates the user's mental health based on the information supplied. This flow creates an efficient pipeline for accurately and reliably predicting mental health using NLP approaches.

5.5 MODULE DESCRIPTIONS

5.5.1 Data Preprocessing

Data preprocessing is the most important step in the NLP pipeline that involves transforming raw data into a structured format appropriate for model training.

This stage includes:

- Text normalization is used to convert all the lowercase, and delete punctuation, special characters, and stop words.
- Tokenization is used to split text into individual tokens using the spacy tokenizer.
- Handling missing values entails replacing or eliminating missing text items.
- Label encoding is used for converting category target labels into a numerical representation.

5.5.2 BERT Pre-trained model

Google created the Bidirectional Encoder Representations from Transformers (BERT) model, which serves as the basic language model. It can grasp a word's context depending on its surrounds (both left and right context).

We take a pre-trained BERT model, such as bert-base-uncased, then fine-tune it for our particular mental health dataset. The model uses deep semantic comprehension and is particularly effective at recognizing emotional subtleties in text.

BERT's main strengths in this context are:

- Identifies sentence-level and contextual links.
- Handles complicated sentence constructions.
- It outperforms typical NLP models in categorization tasks.

5.5.3 Optimization with AdamW

To improve the training process, we employ the AdamW optimizer, which stands for Adaptive Moment Estimation with Weight Decay. It improves on standard Adam by separating weight decay and gradient updating, resulting in higher generalization.

- Advantages of AdamW for fine-tuning:
- prevents overfitting by regulated regularization.
- stabilizes training for transformer models.
- Efficient convergence when training on tiny datasets.

5.5.2 Evaluation Matrix

After training, the model's performance is assessed using conventional classification measures like:

- Accuracy measures the percentage of accurately anticipated instances.
- Precision, recall, and F1-score - Assess the balance of true predictions and erroneous positives/negatives.
- Confusion Matrix - Visualizes model performance across several mental health classes.

5.5.5 Fine-tuned the model

The model is fine-tuned further in response to the assessment outcomes.

This includes:

- Adjusting the learning rate, batch size, and number of epochs
- Freezing or unfreezing particular BERT layers for controlled training.
- Adding dropout and regularization layers

5.5.6 Predicting the mental health

Once trained and fine-tuned, the model may be used to categorize fresh user inputs as mental health problems.

Outputs can include:

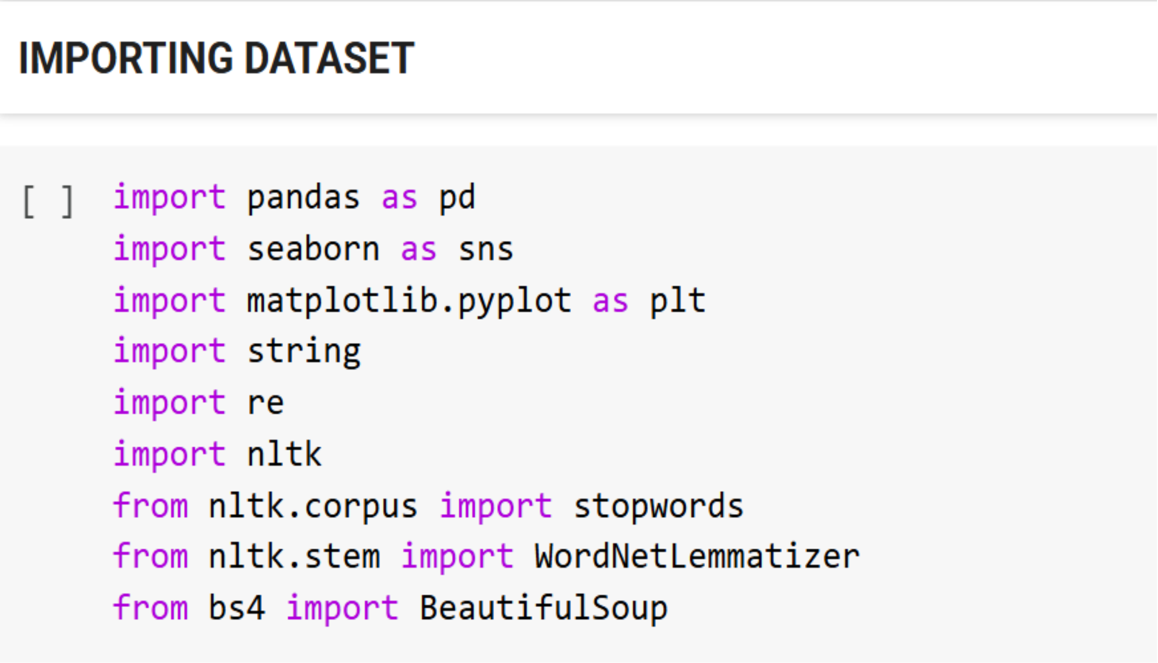
- Predicted class labels (e.g., anxiety, depression, etc.,)
- Class probabilities

Chapter 6

Implementation & Testing

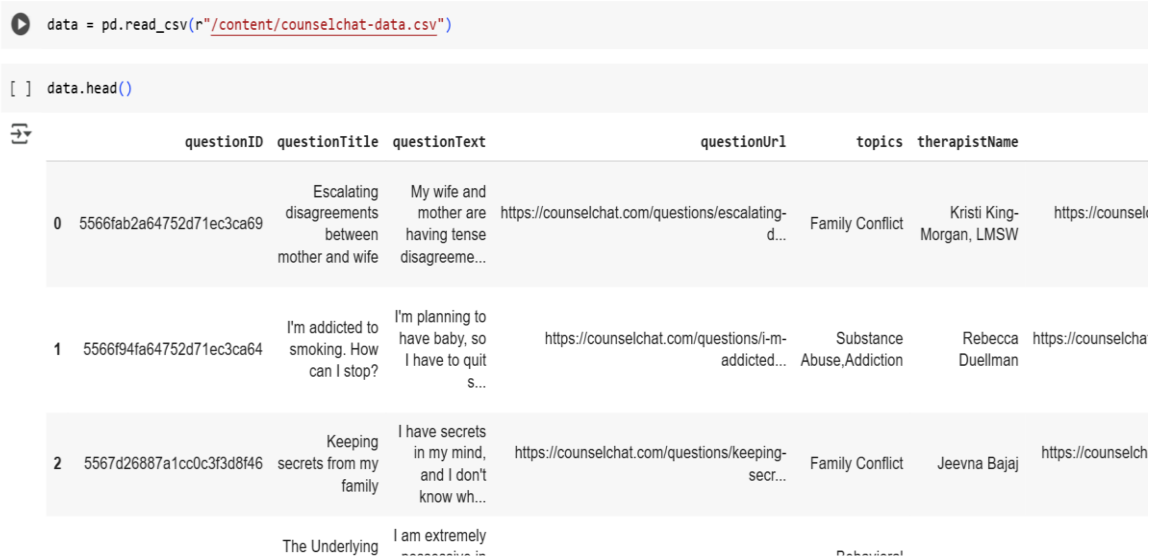
6.1SAMPLE CODE

6.1.1 Importing Dataset



IMPORTING DATASET

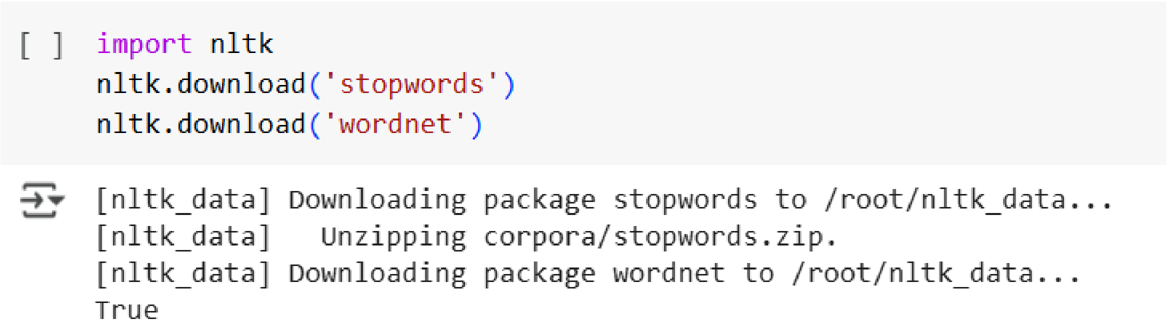
```
[ ]  import pandas as pd
     import seaborn as sns
     import matplotlib.pyplot as plt
     import string
     import re
     import nltk
     from nltk.corpus import stopwords
     from nltk.stem import WordNetLemmatizer
     from bs4 import BeautifulSoup
```



```
data = pd.read_csv(r"/content/counselchat-data.csv")
```

```
[ ]  data.head()
```

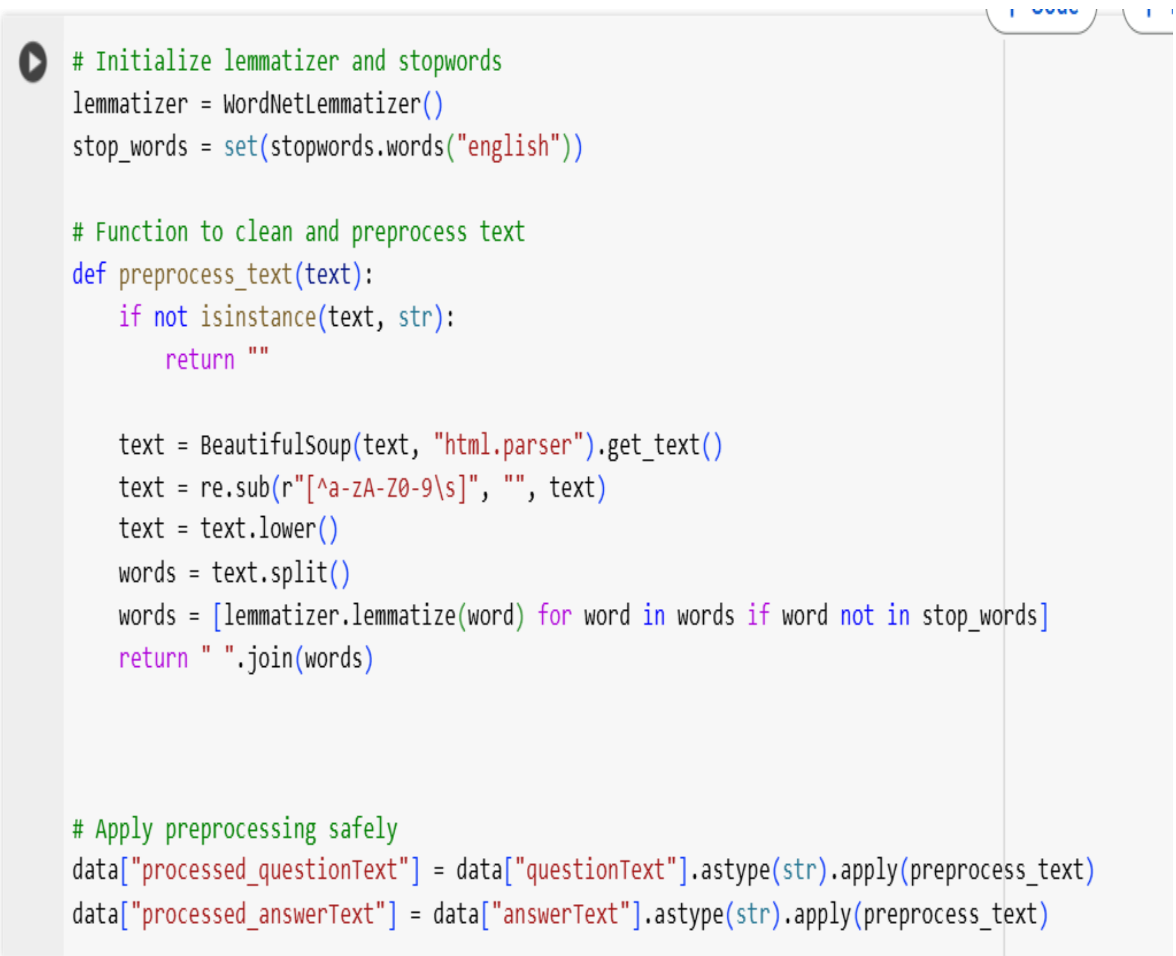
	questionID	questionTitle	questionText	questionUrl	topics	therapistName	
0	5566fab2a64752d71ec3ca69	Escalating disagreements between mother and wife	My wife and mother are having tense disagreeme...	https://counselchat.com/questions/escalating-d...	Family Conflict	Kristi King-Morgan, LMSW	https://counsel
1	5566f94fa64752d71ec3ca64	I'm addicted to smoking. How can I stop?	I'm planning to have baby, so I have to quit s...	https://counselchat.com/questions/i-m-addicted...	Substance Abuse,Addiction	Rebecca Duellman	https://counselcha
2	5567d26887a1cc0c3f3d8f46	Keeping secrets from my family	I have secrets in my mind, and I don't know wh...	https://counselchat.com/questions/keeping-secr...	Family Conflict	Jeevna Bajaj	https://counselch
		The Underlying	I am extremely		Behavioral		



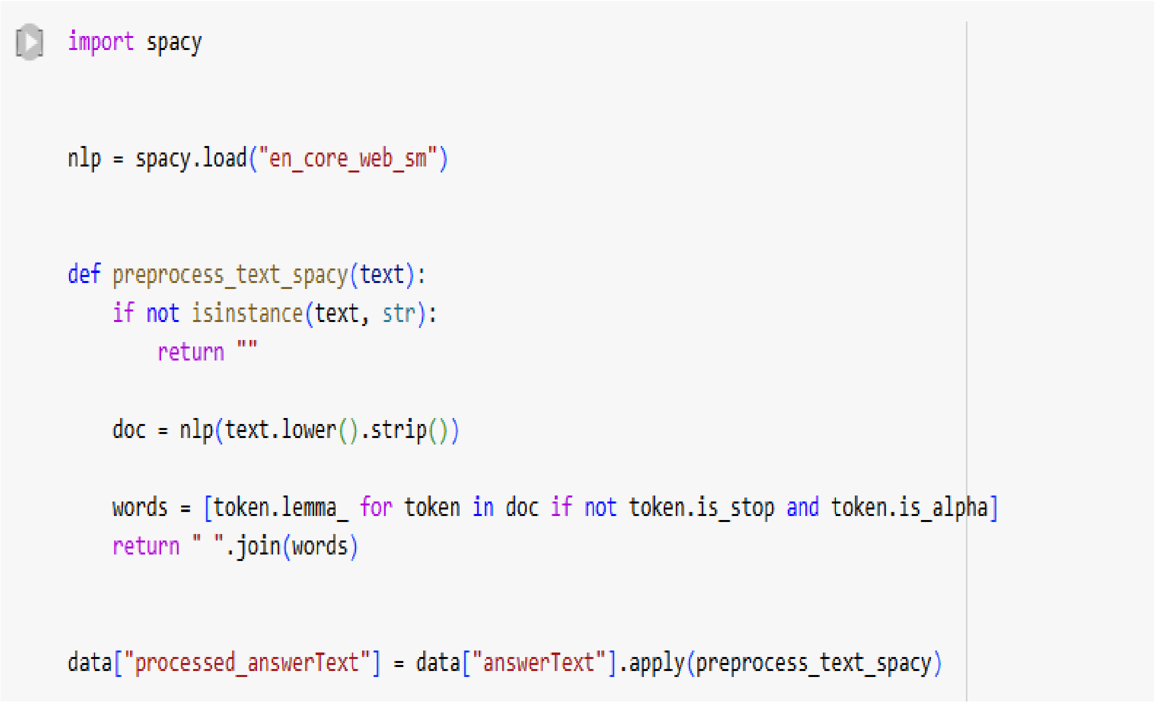
```
[ ]  import nltk
     nltk.download('stopwords')
     nltk.download('wordnet')
```

```
[nltk_data] Downloading package stopwords to /root/nltk_data...
[nltk_data]   Unzipping corpora/stopwords.zip.
[nltk_data] Downloading package wordnet to /root/nltk_data...
True
```

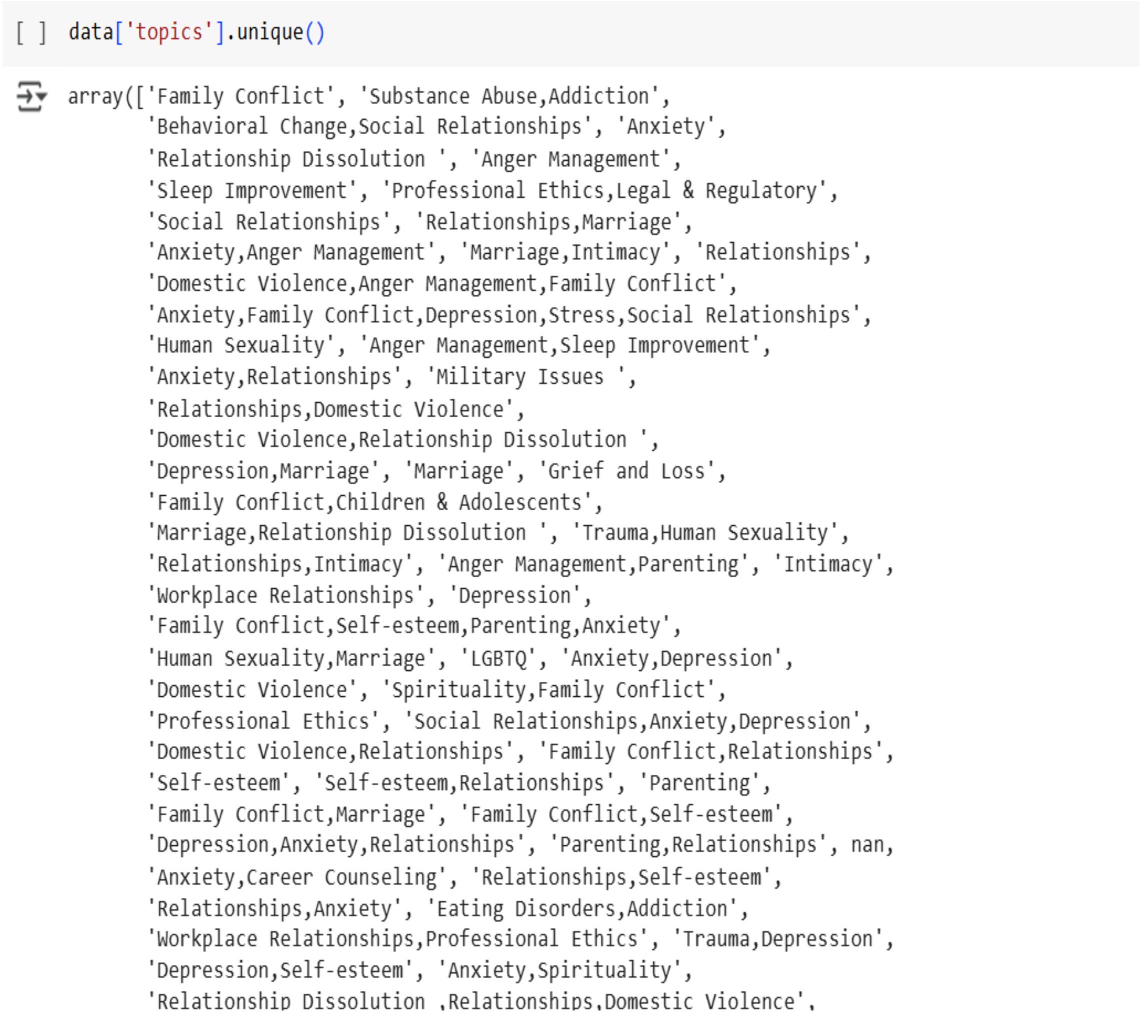
6.1.2 Data Preprocessing



```
# Initialize lemmatizer and stopwords
lemmatizer = WordNetLemmatizer()
stop_words = set(stopwords.words("english"))

# Function to clean and preprocess text
def preprocess_text(text):
    if not isinstance(text, str):
        return ""

    text = BeautifulSoup(text, "html.parser").get_text()
    text = re.sub(r"[^a-zA-Z0-9\s]", "", text)
    text = text.lower()
    words = text.split()
    words = [lemmatizer.lemmatize(word) for word in words if word not in stop_words]
    return " ".join(words)



# Apply preprocessing safely
data["processed_questionText"] = data["questionText"].astype(str).apply(preprocess_text)
data["processed_answerText"] = data["answerText"].astype(str).apply(preprocess_text)
```

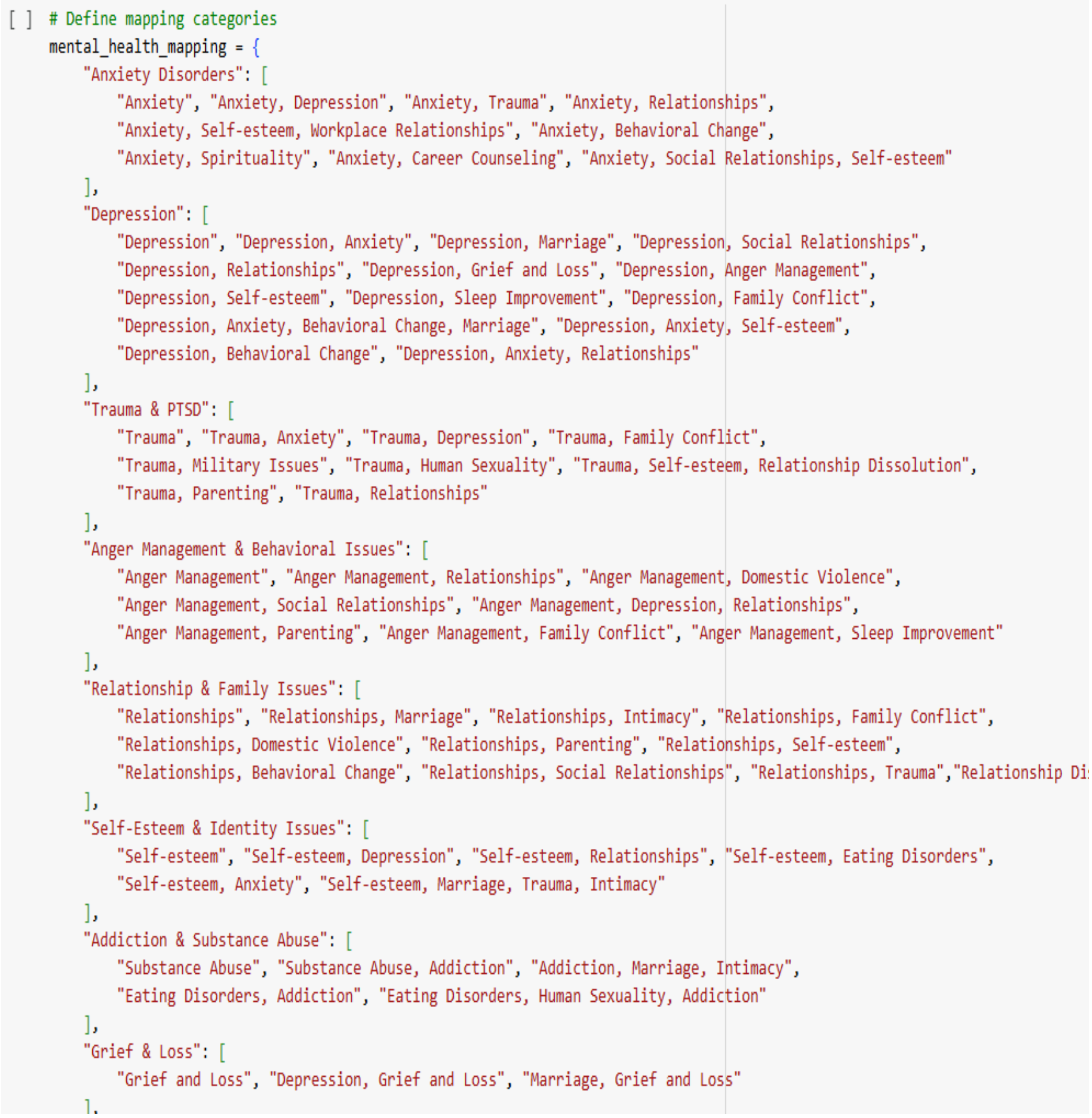


```python
import spacy


nlp = spacy.load("en_core_web_sm")


def preprocess_text_spacy(text):
    if not isinstance(text, str):
        return ""

    doc = nlp(text.lower().strip())

    words = [token.lemma_ for token in doc if not token.is_stop and token.is_alpha]
    return " ".join(words)


data["processed_answerText"] = data["answerText"].apply(preprocess_text_spacy)
```

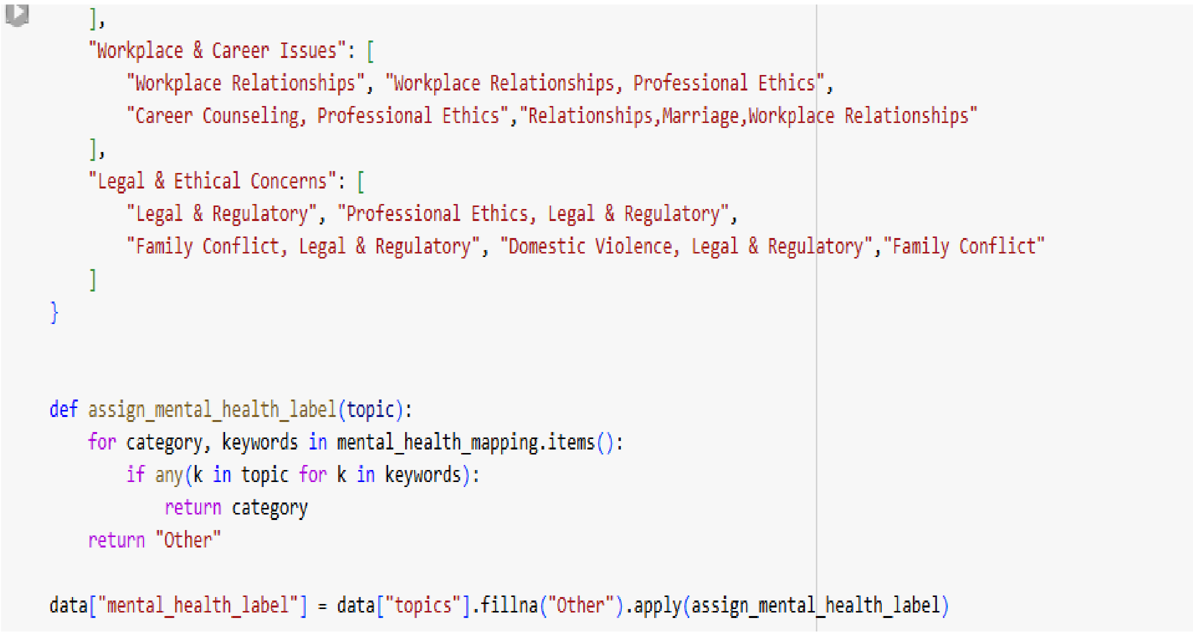


```
[ ]  data['topics'].unique()
```

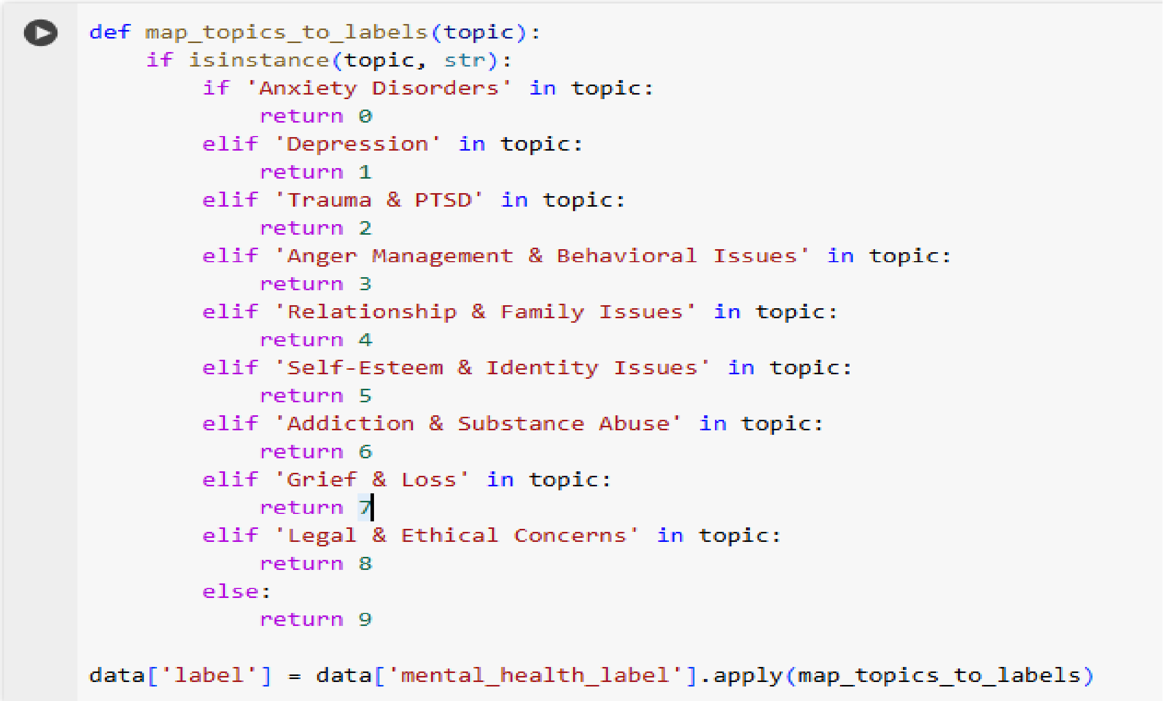
```
array(['Family Conflict', 'Substance Abuse,Addiction',
       'Behavioral Change,Social Relationships', 'Anxiety',
       'Relationship Dissolution ', 'Anger Management',
       'Sleep Improvement', 'Professional Ethics,Legal & Regulatory',
       'Social Relationships', 'Relationships,Marriage',
       'Anxiety,Anger Management', 'Marriage,Intimacy', 'Relationships',
       'Domestic Violence,Anger Management,Family Conflict',
       'Anxiety,Family Conflict,Depression,Stress,Social Relationships',
       'Human Sexuality', 'Anger Management,Sleep Improvement',
       'Anxiety,Relationships', 'Military Issues ',
       'Relationships,Domestic Violence',
       'Domestic Violence,Relationship Dissolution ',
       'Depression,Marriage', 'Marriage', 'Grief and Loss',
       'Family Conflict,Children & Adolescents',
       'Marriage,Relationship Dissolution ', 'Trauma,Human Sexuality',
       'Relationships,Intimacy', 'Anger Management,Parenting', 'Intimacy',
       'Workplace Relationships', 'Depression',
       'Family Conflict,Self-esteem,Parenting,Anxiety',
       'Human Sexuality,Marriage', 'LGBTQ', 'Anxiety,Depression',
       'Domestic Violence', 'Spirituality,Family Conflict',
       'Professional Ethics', 'Social Relationships,Anxiety,Depression',
       'Domestic Violence,Relationships', 'Family Conflict,Relationships',
       'Self-esteem', 'Self-esteem,Relationships', 'Parenting',
       'Family Conflict,Marriage', 'Family Conflict,Self-esteem',
       'Depression,Anxiety,Relationships', 'Parenting,Relationships', nan,
       'Anxiety,Career Counseling', 'Relationships,Self-esteem',
       'Relationships,Anxiety', 'Eating Disorders,Addiction',
       'Workplace Relationships,Professional Ethics', 'Trauma,Depression',
       'Depression,Self-esteem', 'Anxiety,Spirituality',
       'Relationship Dissolution ,Relationships,Domestic Violence',
```



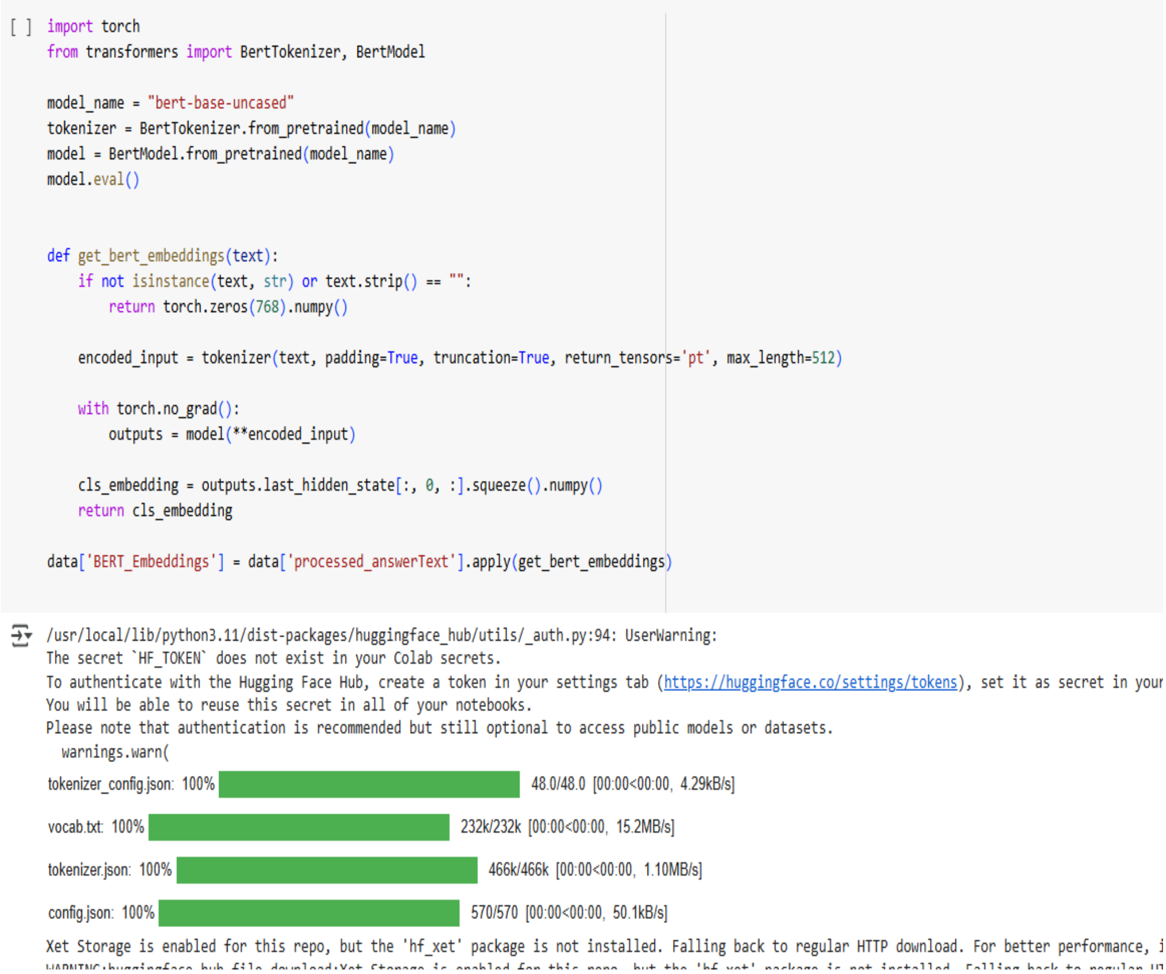
```
[ ]  # Define mapping categories
     mental_health_mapping = {
         "Anxiety Disorders": [
             "Anxiety", "Anxiety, Depression", "Anxiety, Trauma", "Anxiety, Relationships",
             "Anxiety, Self-esteem, Workplace Relationships", "Anxiety, Behavioral Change",
             "Anxiety, Spirituality", "Anxiety, Career Counseling", "Anxiety, Social Relationships, Self-esteem"
         ],
         "Depression": [
             "Depression", "Depression, Anxiety", "Depression, Marriage", "Depression, Social Relationships",
             "Depression, Relationships", "Depression, Grief and Loss", "Depression, Anger Management",
             "Depression, Self-esteem", "Depression, Sleep Improvement", "Depression, Family Conflict",
             "Depression, Anxiety, Behavioral Change, Marriage", "Depression, Anxiety, Self-esteem",
             "Depression, Behavioral Change", "Depression, Anxiety, Relationships"
         ],
         "Trauma & PTSD": [
             "Trauma", "Trauma, Anxiety", "Trauma, Depression", "Trauma, Family Conflict",
             "Trauma, Military Issues", "Trauma, Human Sexuality", "Trauma, Self-esteem, Relationship Dissolution",
             "Trauma, Parenting", "Trauma, Relationships"
         ],
         "Anger Management & Behavioral Issues": [
             "Anger Management", "Anger Management, Relationships", "Anger Management, Domestic Violence",
             "Anger Management, Social Relationships", "Anger Management, Depression, Relationships",
             "Anger Management, Parenting", "Anger Management, Family Conflict", "Anger Management, Sleep Improvement"
         ],
         "Relationship & Family Issues": [
             "Relationships", "Relationships, Marriage", "Relationships, Intimacy", "Relationships, Family Conflict",
             "Relationships, Domestic Violence", "Relationships, Parenting", "Relationships, Self-esteem",
             "Relationships, Behavioral Change", "Relationships, Social Relationships", "Relationships, Trauma","Relationship Di
         ],
         "Self-Esteem & Identity Issues": [
             "Self-esteem", "Self-esteem, Depression", "Self-esteem, Relationships", "Self-esteem, Eating Disorders",
             "Self-esteem, Anxiety", "Self-esteem, Marriage, Trauma, Intimacy"
         ],
         "Addiction & Substance Abuse": [
             "Substance Abuse", "Substance Abuse, Addiction", "Addiction, Marriage, Intimacy",
             "Eating Disorders, Addiction", "Eating Disorders, Human Sexuality, Addiction"
         ],
         "Grief & Loss": [
             "Grief and Loss", "Depression, Grief and Loss", "Marriage, Grief and Loss"
         ],
```



```
         ],
         "Workplace & Career Issues": [
             "Workplace Relationships", "Workplace Relationships, Professional Ethics",
             "Career Counseling, Professional Ethics","Relationships,Marriage,Workplace Relationships"
         ],
         "Legal & Ethical Concerns": [
             "Legal & Regulatory", "Professional Ethics, Legal & Regulatory",
             "Family Conflict, Legal & Regulatory", "Domestic Violence, Legal & Regulatory","Family Conflict"
         ]
     }


     def assign_mental_health_label(topic):
         for category, keywords in mental_health_mapping.items():
             if any(k in topic for k in keywords):
                 return category
         return "Other"

     data["mental_health_label"] = data["topics"].fillna("Other").apply(assign_mental_health_label)
```



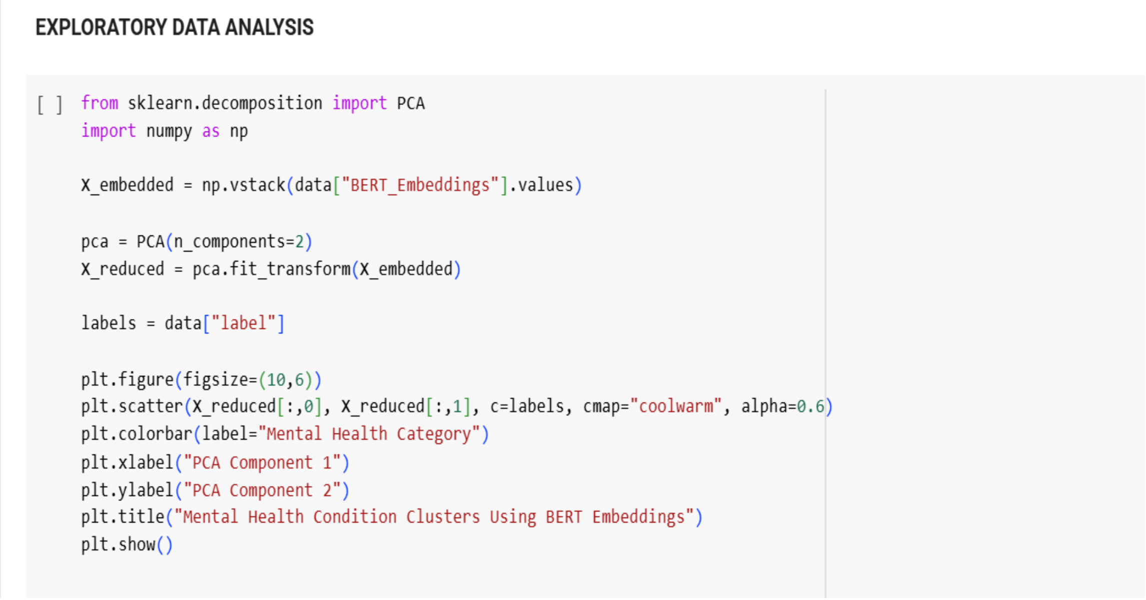
```python
def map_topics_to_labels(topic):
    if isinstance(topic, str):
        if 'Anxiety Disorders' in topic:
            return 0
        elif 'Depression' in topic:
            return 1
        elif 'Trauma & PTSD' in topic:
            return 2
        elif 'Anger Management & Behavioral Issues' in topic:
            return 3
        elif 'Relationship & Family Issues' in topic:
            return 4
        elif 'Self-Esteem & Identity Issues' in topic:
            return 5
        elif 'Addiction & Substance Abuse' in topic:
            return 6
        elif 'Grief & Loss' in topic:
            return 7
        elif 'Legal & Ethical Concerns' in topic:
            return 8
        else:
            return 9

data['label'] = data['mental_health_label'].apply(map_topics_to_labels)
```



```python
import torch
from transformers import BertTokenizer, BertModel

model_name = "bert-base-uncased"
tokenizer = BertTokenizer.from_pretrained(model_name)
model = BertModel.from_pretrained(model_name)
model.eval()


def get_bert_embeddings(text):
    if not isinstance(text, str) or text.strip() == "":
        return torch.zeros(768).numpy()

    encoded_input = tokenizer(text, padding=True, truncation=True, return_tensors='pt', max_length=512)

    with torch.no_grad():
        outputs = model(**encoded_input)

    cls_embedding = outputs.last_hidden_state[:, 0, :].squeeze().numpy()
    return cls_embedding

data['BERT_Embeddings'] = data['processed_answerText'].apply(get_bert_embeddings)
```

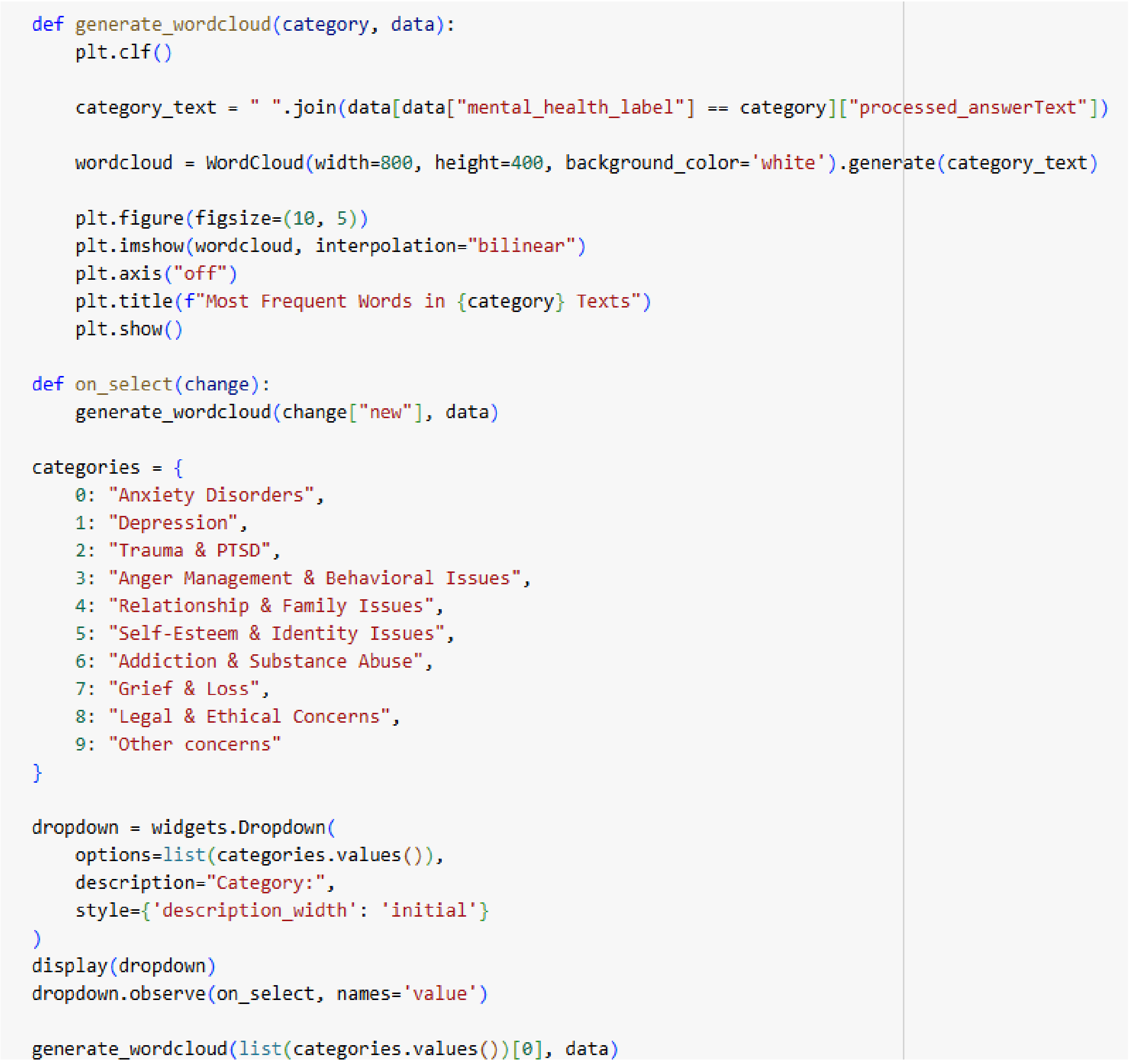
```
/usr/local/lib/python3.11/dist-packages/huggingface_hub/utils/_auth.py:94: UserWarning:
The secret `HF_TOKEN` does not exist in your Colab secrets.
To authenticate with the Hugging Face Hub, create a token in your settings tab (https://huggingface.co/settings/tokens), set it as secret in your
You will be able to reuse this secret in all of your notebooks.
Please note that authentication is recommended but still optional to access public models or datasets.
  warnings.warn(
tokenizer_config.json: 100% ████████████ 48.0/48.0 [00:00<00:00, 4.29kB/s]
vocab.txt: 100% ████████████ 232k/232k [00:00<00:00, 15.2MB/s]
tokenizer.json: 100% ████████████ 466k/466k [00:00<00:00, 1.10MB/s]
config.json: 100% ████████████ 570/570 [00:00<00:00, 50.1kB/s]
Xet Storage is enabled for this repo, but the 'hf_xet' package is not installed. Falling back to regular HTTP download. For better performance, i
WARNING:huggingface_hub.file_download:Xet Storage is enabled for this repo, but the 'hf_xet' package is not installed. Falling back to regular HT
```

6.1.3 Exploratory Data Analysis

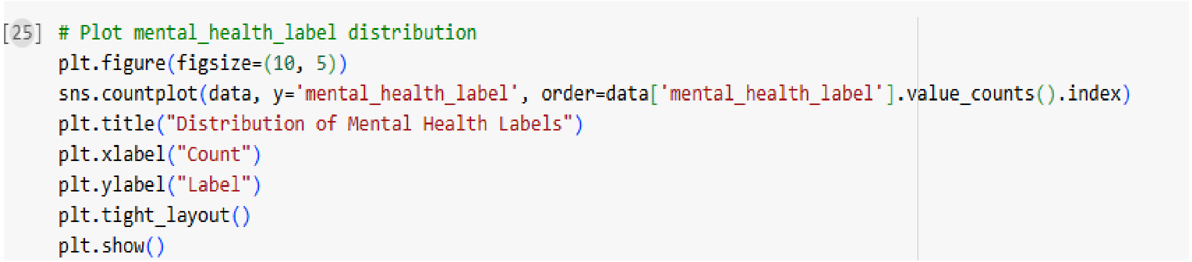


EXPLORATORY DATA ANALYSIS

```
[ ] from sklearn.decomposition import PCA
    import numpy as np

    X_embedded = np.vstack(data["BERT_Embeddings"].values)

    pca = PCA(n_components=2)
    X_reduced = pca.fit_transform(X_embedded)

    labels = data["label"]

    plt.figure(figsize=(10,6))
    plt.scatter(X_reduced[:,0], X_reduced[:,1], c=labels, cmap="coolwarm", alpha=0.6)
    plt.colorbar(label="Mental Health Category")
    plt.xlabel("PCA Component 1")
    plt.ylabel("PCA Component 2")
    plt.title("Mental Health Condition Clusters Using BERT Embeddings")
    plt.show()
```

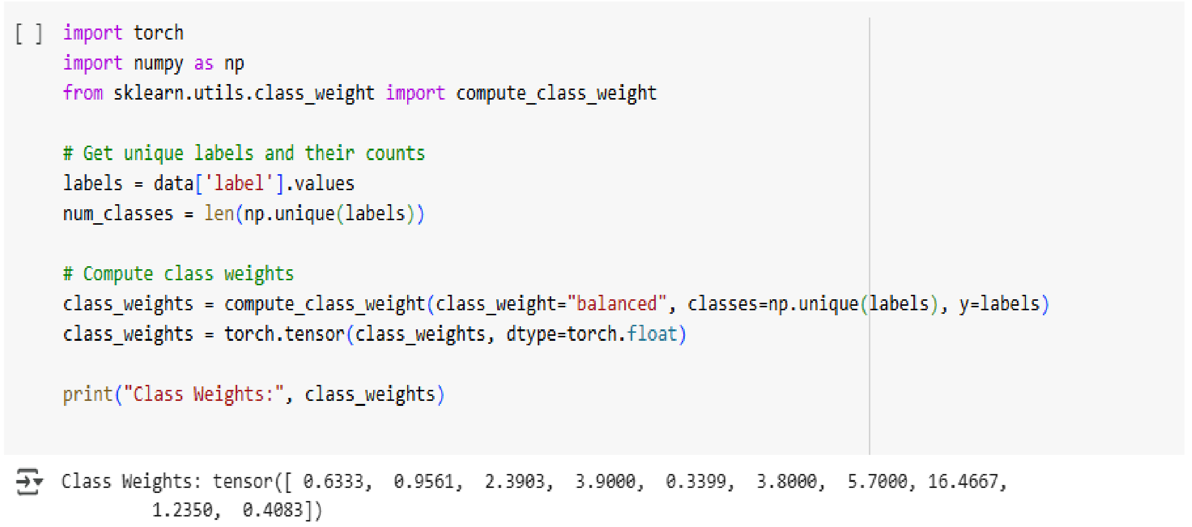


```
def generate_wordcloud(category, data):
    plt.clf()

    category_text = " ".join(data[data["mental_health_label"] == category]["processed_answerText"])

    wordcloud = WordCloud(width=800, height=400, background_color='white').generate(category_text)

    plt.figure(figsize=(10, 5))
    plt.imshow(wordcloud, interpolation="bilinear")
    plt.axis("off")
    plt.title(f"Most Frequent Words in {category} Texts")
    plt.show()

def on_select(change):
    generate_wordcloud(change["new"], data)

categories = {
    0: "Anxiety Disorders",
    1: "Depression",
    2: "Trauma & PTSD",
    3: "Anger Management & Behavioral Issues",
    4: "Relationship & Family Issues",
    5: "Self-Esteem & Identity Issues",
    6: "Addiction & Substance Abuse",
    7: "Grief & Loss",
    8: "Legal & Ethical Concerns",
    9: "Other concerns"
}

dropdown = widgets.Dropdown(
    options=list(categories.values()),
    description="Category:",
    style={'description_width': 'initial'}
)
display(dropdown)
dropdown.observe(on_select, names='value')

generate_wordcloud(list(categories.values())[0], data)
```



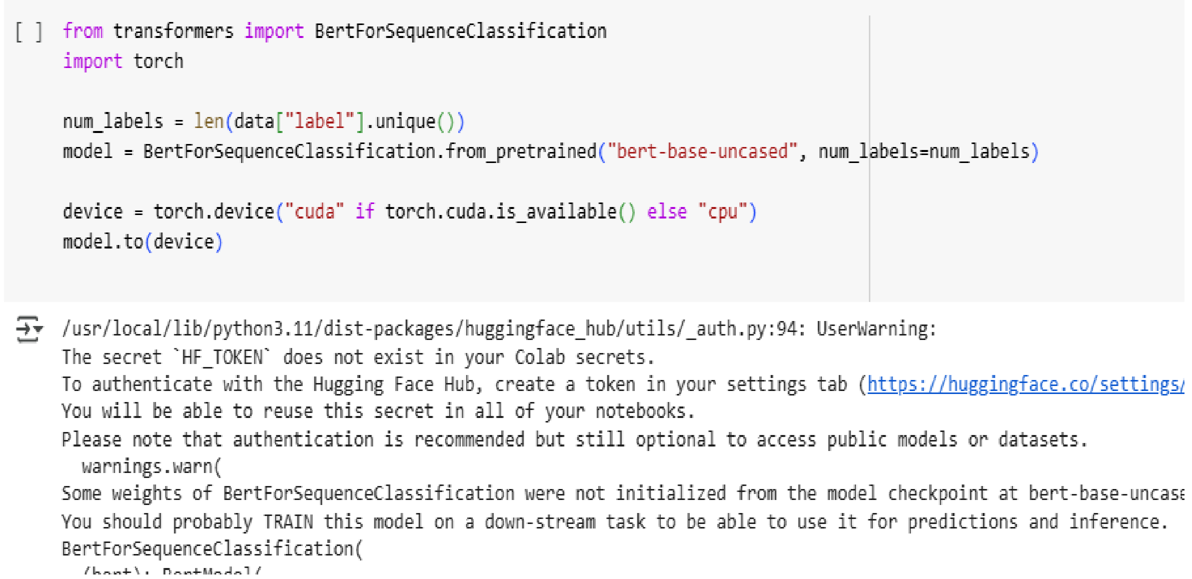
```
[25] # Plot mental_health_label distribution
     plt.figure(figsize=(10, 5))
     sns.countplot(data, y='mental_health_label', order=data['mental_health_label'].value_counts().index)
     plt.title("Distribution of Mental Health Labels")
     plt.xlabel("Count")
     plt.ylabel("Label")
     plt.tight_layout()
     plt.show()
```

6.1.4 BERT pre-trained model



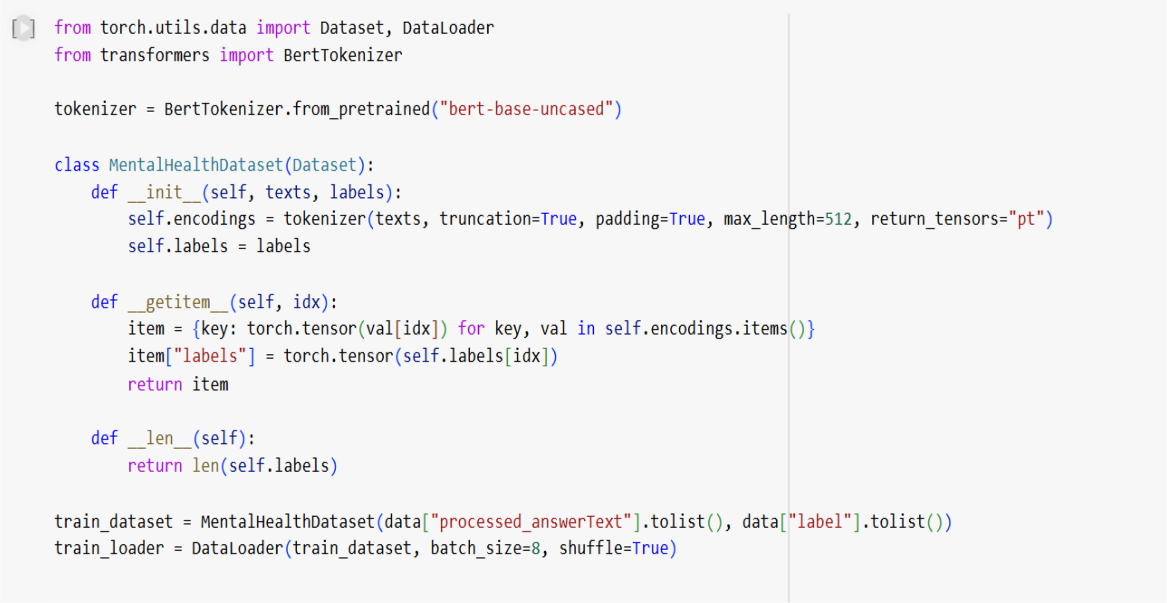
```
[ ]  import torch
     import numpy as np
     from sklearn.utils.class_weight import compute_class_weight

     # Get unique labels and their counts
     labels = data['label'].values
     num_classes = len(np.unique(labels))

     # Compute class weights
     class_weights = compute_class_weight(class_weight="balanced", classes=np.unique(labels), y=labels)
     class_weights = torch.tensor(class_weights, dtype=torch.float)

     print("Class Weights:", class_weights)
```

```
Class Weights: tensor([ 0.6333,  0.9561,  2.3903,  3.9000,  0.3399,  3.8000,  5.7000, 16.4667,
         1.2350,  0.4083])
```

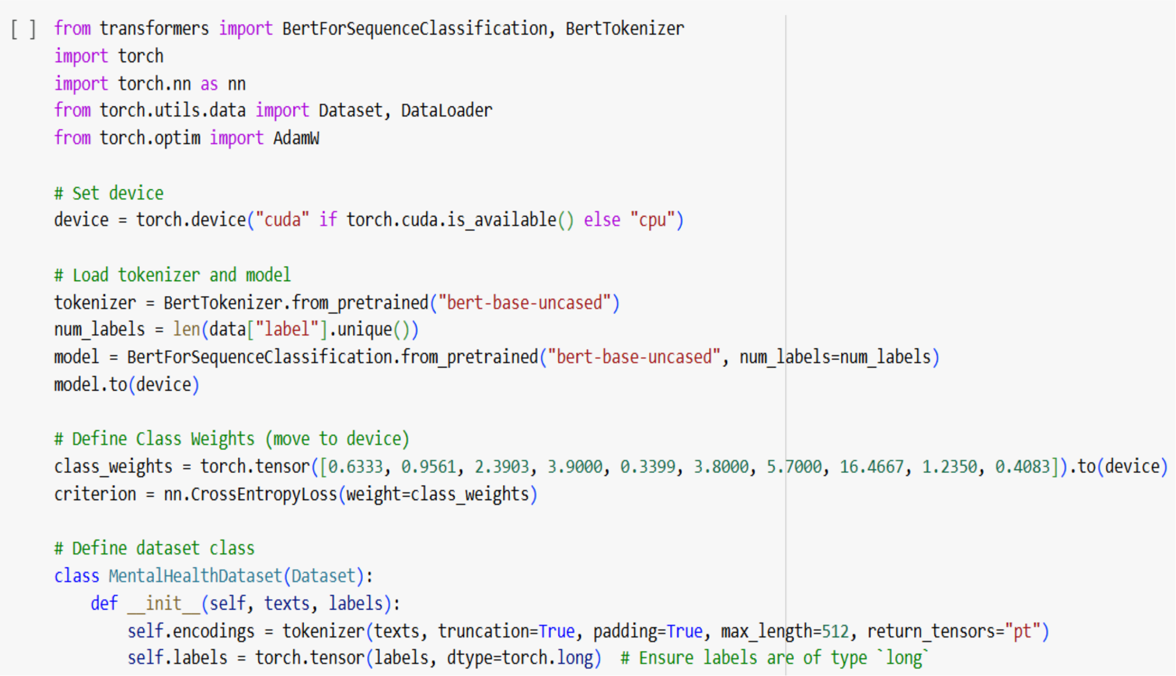


```
[ ]  from transformers import BertForSequenceClassification
     import torch

     num_labels = len(data["label"].unique())
     model = BertForSequenceClassification.from_pretrained("bert-base-uncased", num_labels=num_labels)

     device = torch.device("cuda" if torch.cuda.is_available() else "cpu")
     model.to(device)
```

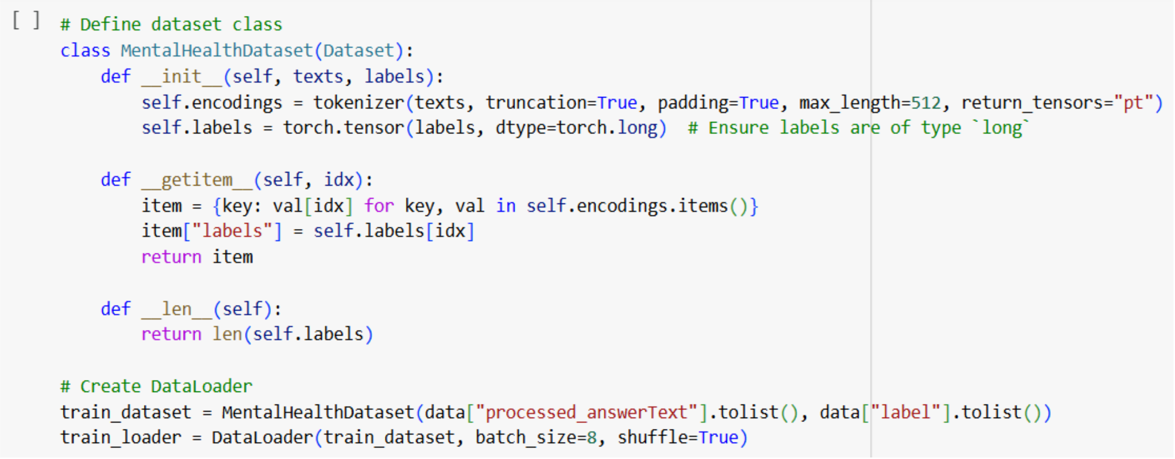
```
/usr/local/lib/python3.11/dist-packages/huggingface_hub/utils/_auth.py:94: UserWarning:
The secret `HF_TOKEN` does not exist in your Colab secrets.
To authenticate with the Hugging Face Hub, create a token in your settings tab (https://huggingface.co/settings/
You will be able to reuse this secret in all of your notebooks.
Please note that authentication is recommended but still optional to access public models or datasets.
  warnings.warn(
Some weights of BertForSequenceClassification were not initialized from the model checkpoint at bert-base-uncase
You should probably TRAIN this model on a down-stream task to be able to use it for predictions and inference.
BertForSequenceClassification(
  (bert): BertModel(
```



```python
from torch.utils.data import Dataset, DataLoader
from transformers import BertTokenizer

tokenizer = BertTokenizer.from_pretrained("bert-base-uncased")

class MentalHealthDataset(Dataset):
    def __init__(self, texts, labels):
        self.encodings = tokenizer(texts, truncation=True, padding=True, max_length=512, return_tensors="pt")
        self.labels = labels

    def __getitem__(self, idx):
        item = {key: torch.tensor(val[idx]) for key, val in self.encodings.items()}
        item["labels"] = torch.tensor(self.labels[idx])
        return item

    def __len__(self):
        return len(self.labels)

train_dataset = MentalHealthDataset(data["processed_answerText"].tolist(), data["label"].tolist())
train_loader = DataLoader(train_dataset, batch_size=8, shuffle=True)
```

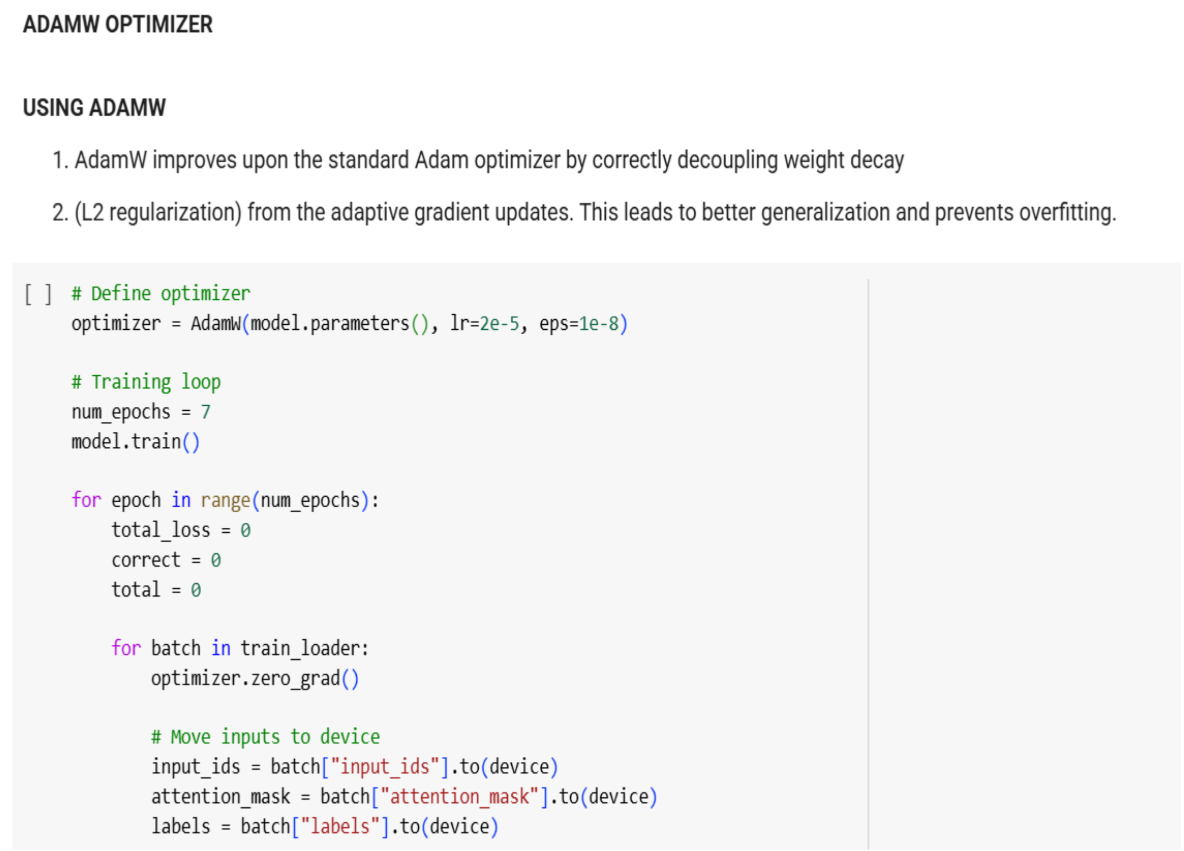


```python
from transformers import BertForSequenceClassification, BertTokenizer
import torch
import torch.nn as nn
from torch.utils.data import Dataset, DataLoader
from torch.optim import AdamW

# Set device
device = torch.device("cuda" if torch.cuda.is_available() else "cpu")

# Load tokenizer and model
tokenizer = BertTokenizer.from_pretrained("bert-base-uncased")
num_labels = len(data["label"].unique())
model = BertForSequenceClassification.from_pretrained("bert-base-uncased", num_labels=num_labels)
model.to(device)

# Define Class Weights (move to device)
class_weights = torch.tensor([0.6333, 0.9561, 2.3903, 3.9000, 0.3399, 3.8000, 5.7000, 16.4667, 1.2350, 0.4083]).to(device)
criterion = nn.CrossEntropyLoss(weight=class_weights)

# Define dataset class
class MentalHealthDataset(Dataset):
    def __init__(self, texts, labels):
        self.encodings = tokenizer(texts, truncation=True, padding=True, max_length=512, return_tensors="pt")
        self.labels = torch.tensor(labels, dtype=torch.long)  # Ensure labels are of type `long`
```



```python
# Define dataset class
class MentalHealthDataset(Dataset):
    def __init__(self, texts, labels):
        self.encodings = tokenizer(texts, truncation=True, padding=True, max_length=512, return_tensors="pt")
        self.labels = torch.tensor(labels, dtype=torch.long)  # Ensure labels are of type `long`

    def __getitem__(self, idx):
        item = {key: val[idx] for key, val in self.encodings.items()}
        item["labels"] = self.labels[idx]
        return item

    def __len__(self):
        return len(self.labels)

# Create DataLoader
train_dataset = MentalHealthDataset(data["processed_answerText"].tolist(), data["label"].tolist())
train_loader = DataLoader(train_dataset, batch_size=8, shuffle=True)
```

6.1.5 ADAMW optimizer

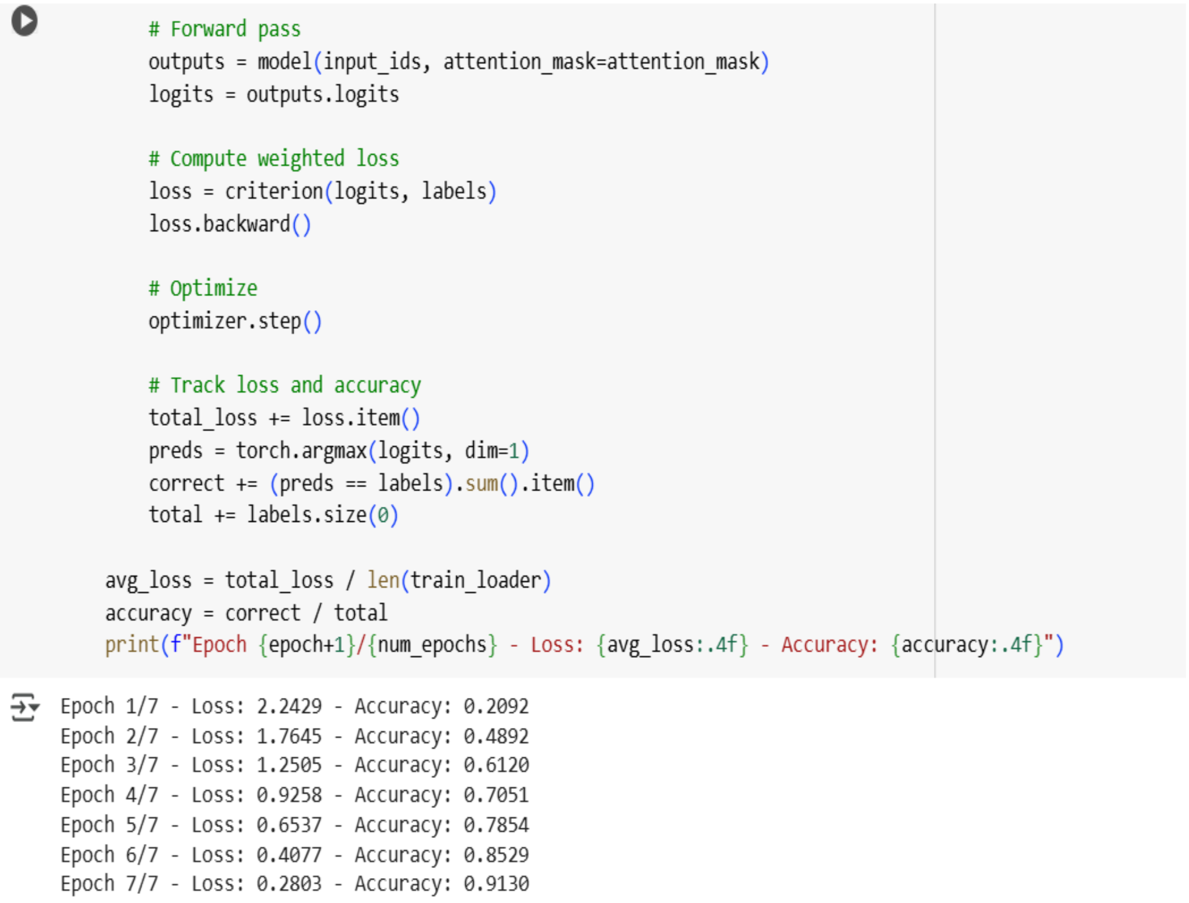


ADAMW OPTIMIZER

USING ADAMW

1. AdamW improves upon the standard Adam optimizer by correctly decoupling weight decay
2. (L2 regularization) from the adaptive gradient updates. This leads to better generalization and prevents overfitting.

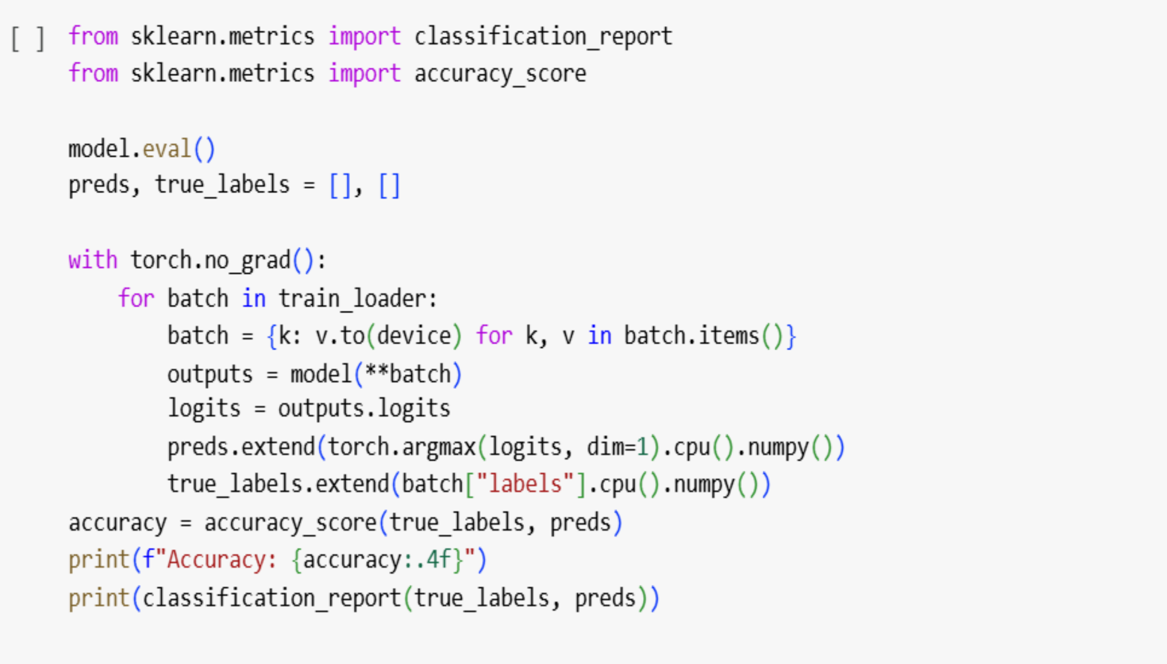
```
[ ] # Define optimizer
    optimizer = AdamW(model.parameters(), lr=2e-5, eps=1e-8)

    # Training loop
    num_epochs = 7
    model.train()

    for epoch in range(num_epochs):
        total_loss = 0
        correct = 0
        total = 0

        for batch in train_loader:
            optimizer.zero_grad()

            # Move inputs to device
            input_ids = batch["input_ids"].to(device)
            attention_mask = batch["attention_mask"].to(device)
            labels = batch["labels"].to(device)
```



```
            # Forward pass
            outputs = model(input_ids, attention_mask=attention_mask)
            logits = outputs.logits

            # Compute weighted loss
            loss = criterion(logits, labels)
            loss.backward()

            # Optimize
            optimizer.step()

            # Track loss and accuracy
            total_loss += loss.item()
            preds = torch.argmax(logits, dim=1)
            correct += (preds == labels).sum().item()
            total += labels.size(0)

        avg_loss = total_loss / len(train_loader)
        accuracy = correct / total
        print(f"Epoch {epoch+1}/{num_epochs} - Loss: {avg_loss:.4f} - Accuracy: {accuracy:.4f}")
```

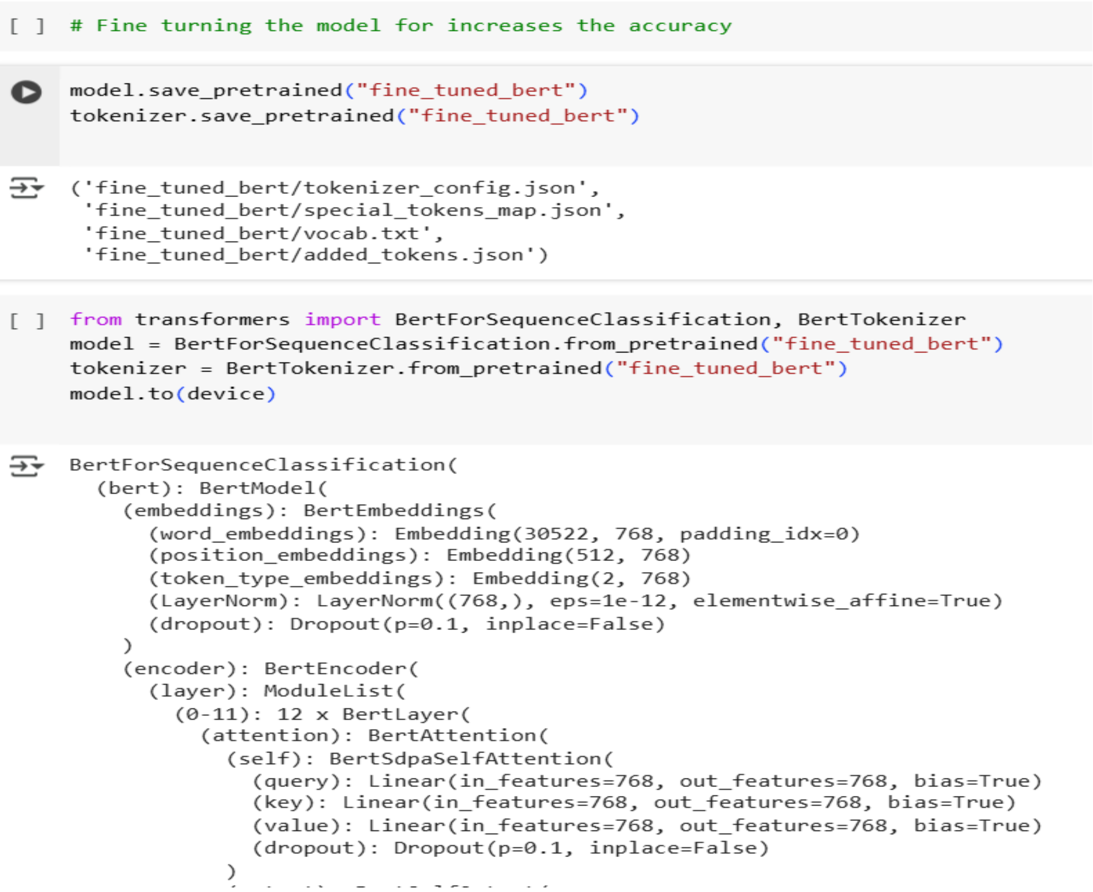
```
Epoch 1/7 - Loss: 2.2429 - Accuracy: 0.2092
Epoch 2/7 - Loss: 1.7645 - Accuracy: 0.4892
Epoch 3/7 - Loss: 1.2505 - Accuracy: 0.6120
Epoch 4/7 - Loss: 0.9258 - Accuracy: 0.7051
Epoch 5/7 - Loss: 0.6537 - Accuracy: 0.7854
Epoch 6/7 - Loss: 0.4077 - Accuracy: 0.8529
Epoch 7/7 - Loss: 0.2803 - Accuracy: 0.9130
```

6.1.6 Evaluation Matrix



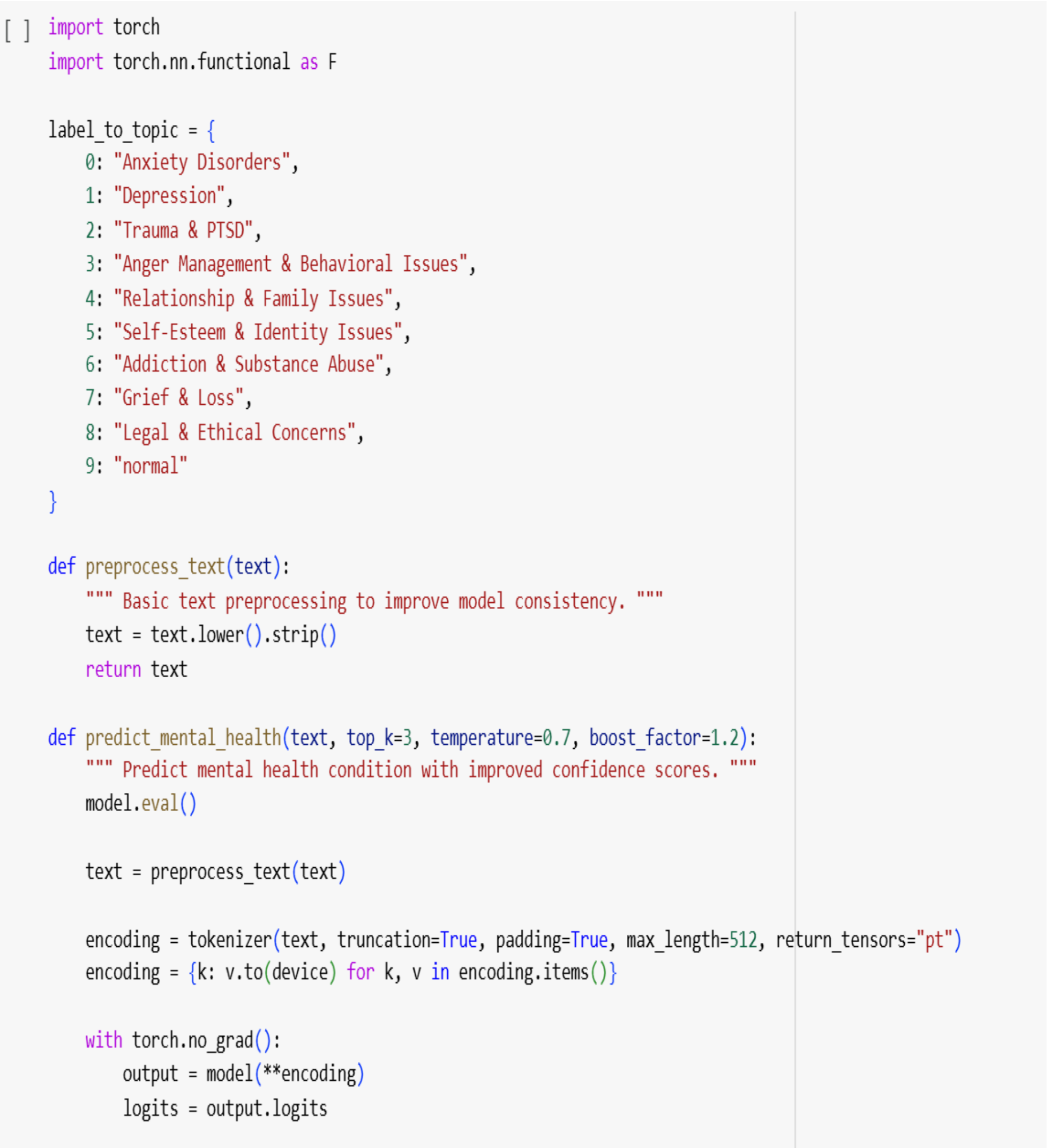
```
[ ] from sklearn.metrics import classification_report
    from sklearn.metrics import accuracy_score

    model.eval()
    preds, true_labels = [], []

    with torch.no_grad():
        for batch in train_loader:
            batch = {k: v.to(device) for k, v in batch.items()}
            outputs = model(**batch)
            logits = outputs.logits
            preds.extend(torch.argmax(logits, dim=1).cpu().numpy())
            true_labels.extend(batch["labels"].cpu().numpy())
    accuracy = accuracy_score(true_labels, preds)
    print(f"Accuracy: {accuracy:.4f}")
    print(classification_report(true_labels, preds))
```

6.2.7 Fine turning model

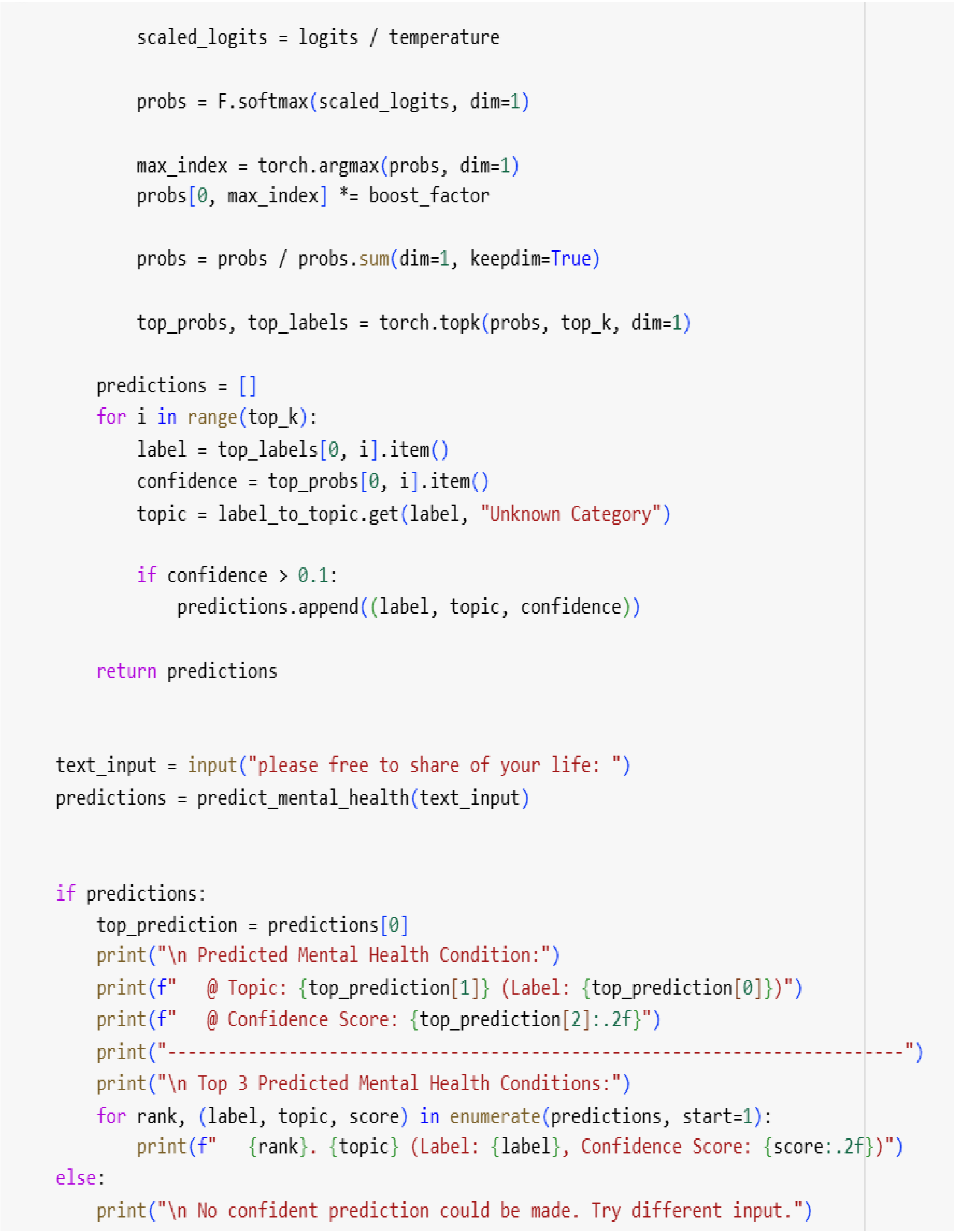


```
[ ] # Fine turning the model for increases the accuracy

    model.save_pretrained("fine_tuned_bert")
    tokenizer.save_pretrained("fine_tuned_bert")

    ('fine_tuned_bert/tokenizer_config.json',
     'fine_tuned_bert/special_tokens_map.json',
     'fine_tuned_bert/vocab.txt',
     'fine_tuned_bert/added_tokens.json')

[ ] from transformers import BertForSequenceClassification, BertTokenizer
    model = BertForSequenceClassification.from_pretrained("fine_tuned_bert")
    tokenizer = BertTokenizer.from_pretrained("fine_tuned_bert")
    model.to(device)

    BertForSequenceClassification(
      (bert): BertModel(
        (embeddings): BertEmbeddings(
          (word_embeddings): Embedding(30522, 768, padding_idx=0)
          (position_embeddings): Embedding(512, 768)
          (token_type_embeddings): Embedding(2, 768)
          (LayerNorm): LayerNorm((768,), eps=1e-12, elementwise_affine=True)
          (dropout): Dropout(p=0.1, inplace=False)
        )
        (encoder): BertEncoder(
          (layer): ModuleList(
            (0-11): 12 x BertLayer(
              (attention): BertAttention(
                (self): BertSdpaSelfAttention(
                  (query): Linear(in_features=768, out_features=768, bias=True)
                  (key): Linear(in_features=768, out_features=768, bias=True)
                  (value): Linear(in_features=768, out_features=768, bias=True)
                  (dropout): Dropout(p=0.1, inplace=False)
                )
```

6.1.7 Prediction for mental health



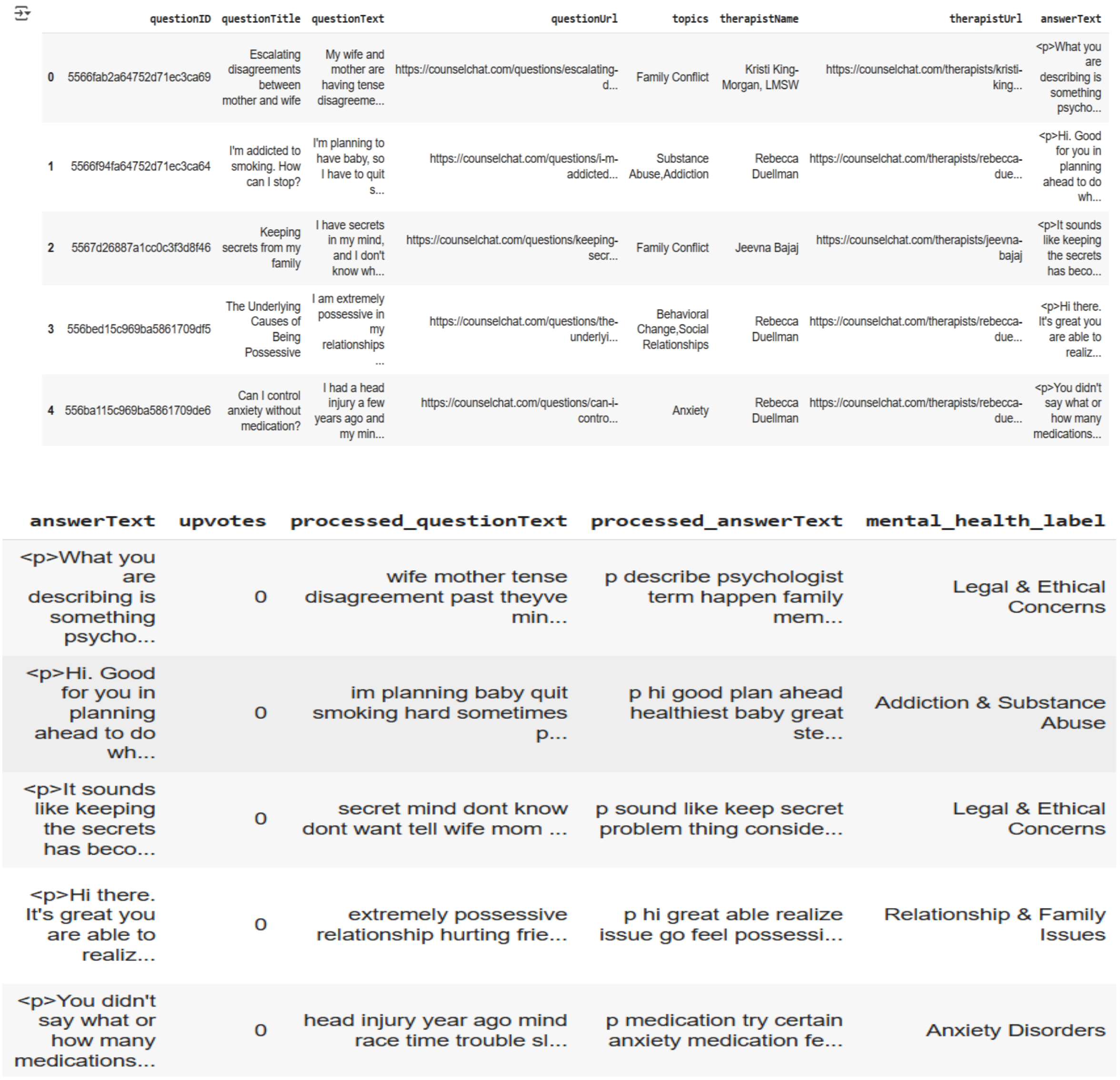
```
import torch
import torch.nn.functional as F

label_to_topic = {
    0: "Anxiety Disorders",
    1: "Depression",
    2: "Trauma & PTSD",
    3: "Anger Management & Behavioral Issues",
    4: "Relationship & Family Issues",
    5: "Self-Esteem & Identity Issues",
    6: "Addiction & Substance Abuse",
    7: "Grief & Loss",
    8: "Legal & Ethical Concerns",
    9: "normal"
}

def preprocess_text(text):
    """ Basic text preprocessing to improve model consistency. """
    text = text.lower().strip()
    return text

def predict_mental_health(text, top_k=3, temperature=0.7, boost_factor=1.2):
    """ Predict mental health condition with improved confidence scores. """
    model.eval()

    text = preprocess_text(text)

    encoding = tokenizer(text, truncation=True, padding=True, max_length=512, return_tensors="pt")
    encoding = {k: v.to(device) for k, v in encoding.items()}

    with torch.no_grad():
        output = model(**encoding)
        logits = output.logits
```



```
        scaled_logits = logits / temperature

        probs = F.softmax(scaled_logits, dim=1)

        max_index = torch.argmax(probs, dim=1)
        probs[0, max_index] *= boost_factor

        probs = probs / probs.sum(dim=1, keepdim=True)

        top_probs, top_labels = torch.topk(probs, top_k, dim=1)

    predictions = []
    for i in range(top_k):
        label = top_labels[0, i].item()
        confidence = top_probs[0, i].item()
        topic = label_to_topic.get(label, "Unknown Category")

        if confidence > 0.1:
            predictions.append((label, topic, confidence))

    return predictions


text_input = input("please free to share of your life: ")
predictions = predict_mental_health(text_input)


if predictions:
    top_prediction = predictions[0]
    print("\n Predicted Mental Health Condition:")
    print(f"   @ Topic: {top_prediction[1]} (Label: {top_prediction[0]})")
    print(f"   @ Confidence Score: {top_prediction[2]:.2f}")
    print("-------------------------------------------------------------------------")
    print("\n Top 3 Predicted Mental Health Conditions:")
    for rank, (label, topic, score) in enumerate(predictions, start=1):
        print(f"   {rank}. {topic} (Label: {label}, Confidence Score: {score:.2f})")
else:
    print("\n No confident prediction could be made. Try different input.")
```

6.2 SAMPLE OUTPUT

6.2.1 Preprocessed data



	questionID	questionTitle	questionText	questionUrl	topics	therapistName	therapistUrl	answerText
0	5566fab2a64752d71ec3ca69	Escalating disagreements between mother and wife	My wife and mother are having tense disagreeme...	https://counselchat.com/questions/escalating-d...	Family Conflict	Kristi King-Morgan, LMSW	https://counselchat.com/therapists/kristi-king...	<p>What you are describing is something psycho...
1	5566f94fa64752d71ec3ca64	I'm addicted to smoking. How can I stop?	I'm planning to have baby, so I have to quit s...	https://counselchat.com/questions/i-m-addicted...	Substance Abuse,Addiction	Rebecca Duellman	https://counselchat.com/therapists/rebecca-due...	<p>Hi. Good for you in planning ahead to do wh...
2	5567d26887a1cc0c3f3d8f46	Keeping secrets from my family	I have secrets in my mind, and I don't know wh...	https://counselchat.com/questions/keeping-secr...	Family Conflict	Jeevna Bajaj	https://counselchat.com/therapists/jeevna-bajaj	<p>It sounds like keeping the secrets has beco...
3	556bed15c969ba5861709df5	The Underlying Causes of Being Possessive	I am extremely possessive in my relationships ...	https://counselchat.com/questions/the-underlyi...	Behavioral Change,Social Relationships	Rebecca Duellman	https://counselchat.com/therapists/rebecca-due...	<p>Hi there. It's great you are able to realiz...
4	556ba115c969ba5861709de6	Can I control anxiety without medication?	I had a head injury a few years ago and my min...	https://counselchat.com/questions/can-i-contro...	Anxiety	Rebecca Duellman	https://counselchat.com/therapists/rebecca-due...	<p>You didn't say what or how many medications...

answerText	upvotes	processed_questionText	processed_answerText	mental_health_label
<p>What you are describing is something psycho...	0	wife mother tense disagreement past theyve min...	p describe psychologist term happen family mem...	Legal & Ethical Concerns
<p>Hi. Good for you in planning ahead to do wh...	0	im planning baby quit smoking hard sometimes p...	p hi good plan ahead healthiest baby great ste...	Addiction & Substance Abuse
<p>It sounds like keeping the secrets has beco...	0	secret mind dont know dont want tell wife mom ...	p sound like keep secret problem thing conside...	Legal & Ethical Concerns
<p>Hi there. It's great you are able to realiz...	0	extremely possessive relationship hurting frie...	p hi great able realize issue go feel possessi...	Relationship & Family Issues
<p>You didn't say what or how many medications...	0	head injury year ago mind race time trouble sl...	p medication try certain anxiety medication fe...	Anxiety Disorders

fig 6.1 pre-processed data

6.2.2 Insights from dataset

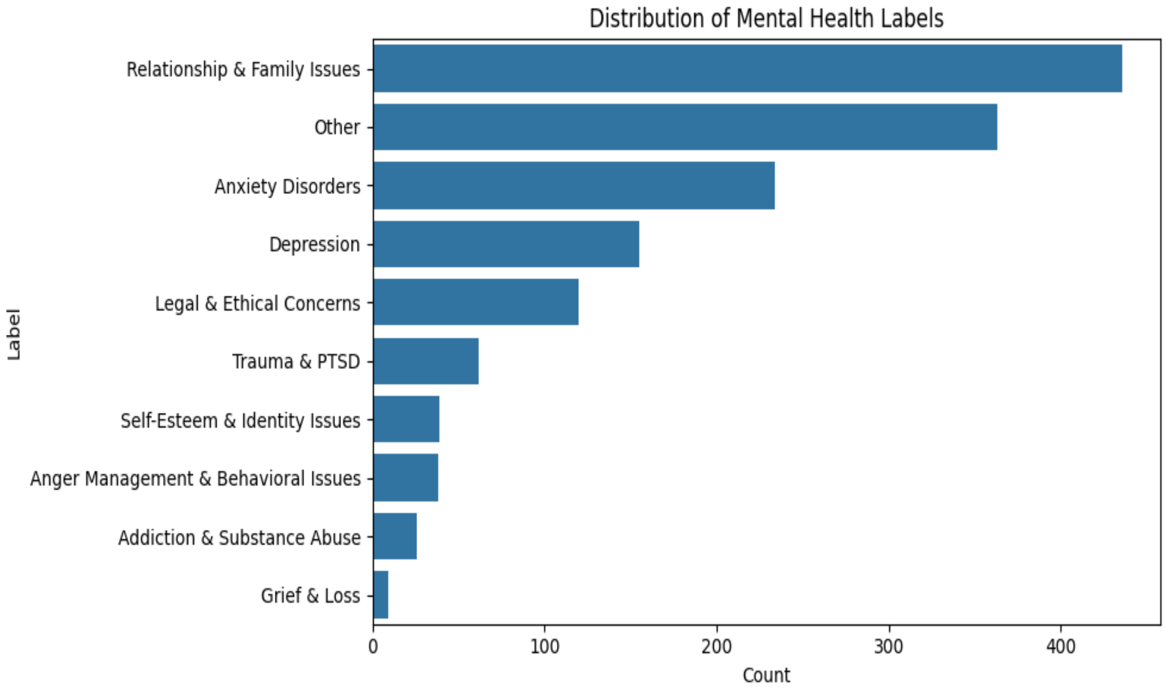


fig 6.2 distribution of mental health label

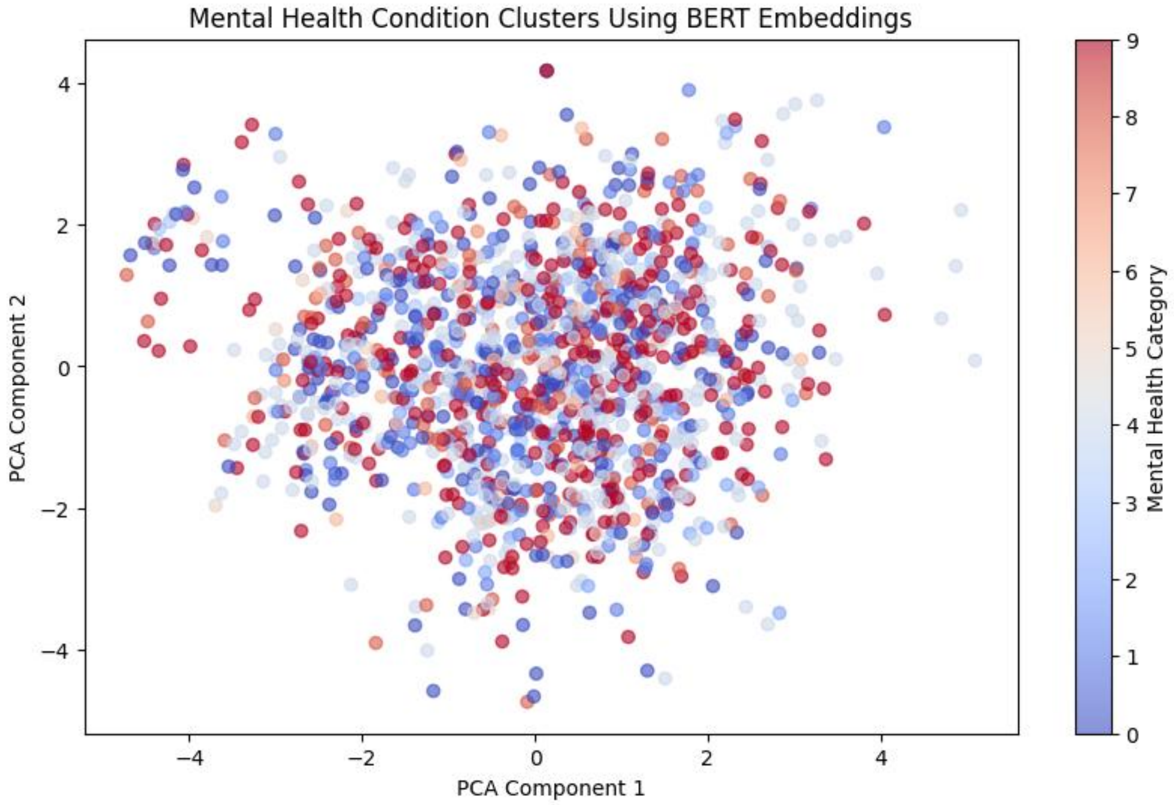


fig 6.3 principal component analysis

6.2.3 Word cloud of all categories

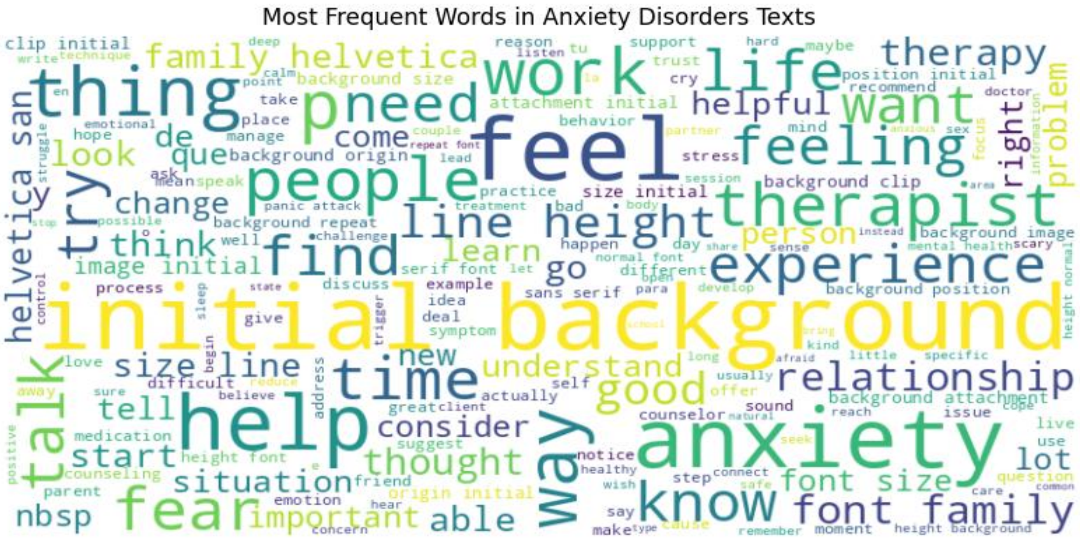


fig 6.4 word cloud for anxiety disorders text

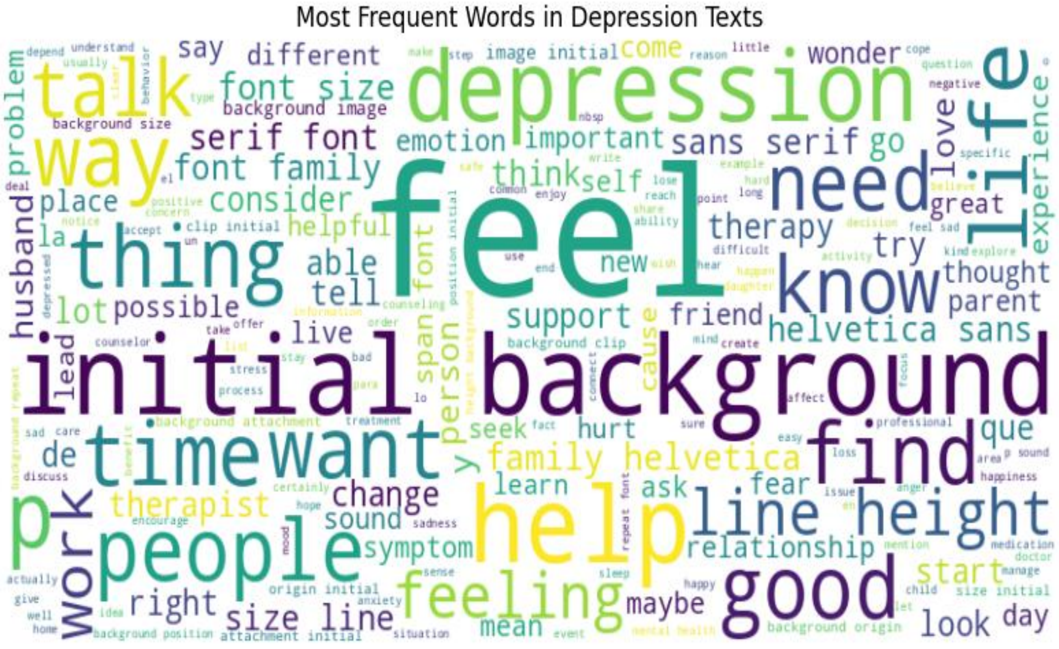


fig 6.5 word cloud for depression texts

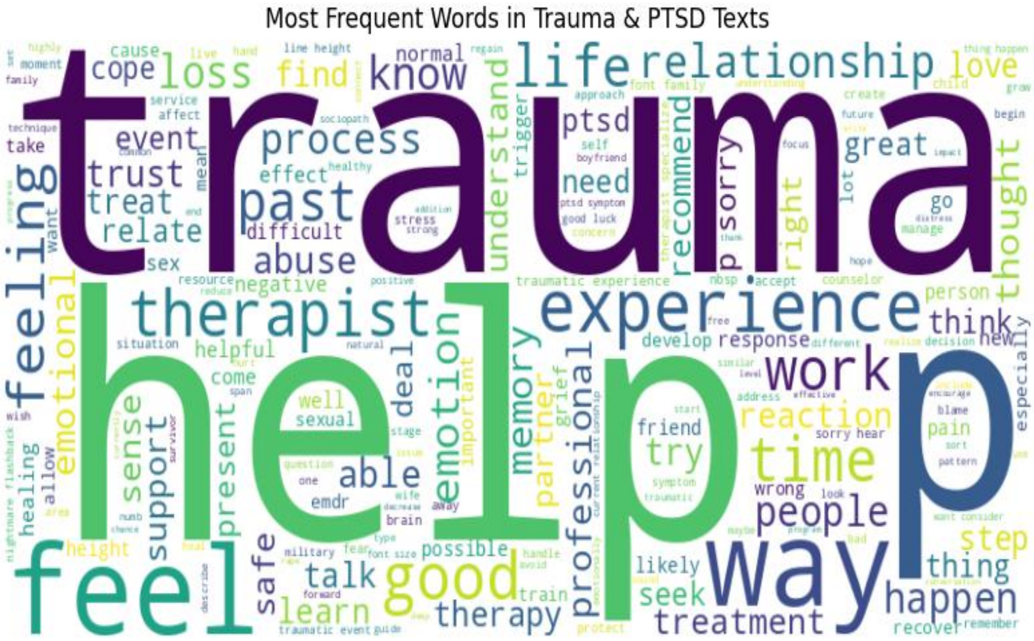


fig 6.6 word cloud for Trauma & PTSD texts

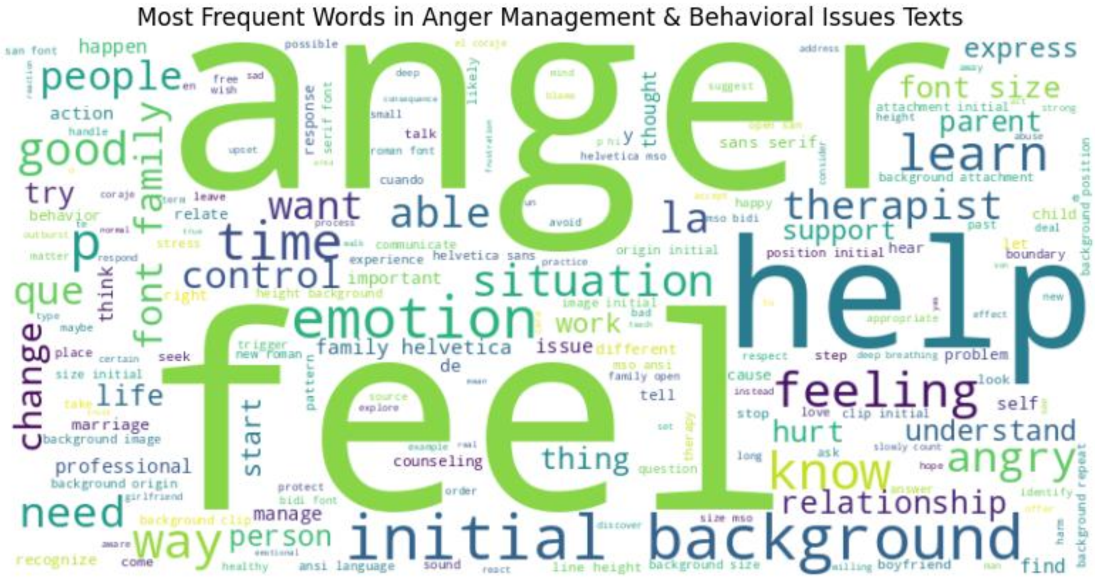


fig 6.7 word cloud for anger management & behavioral issues texts

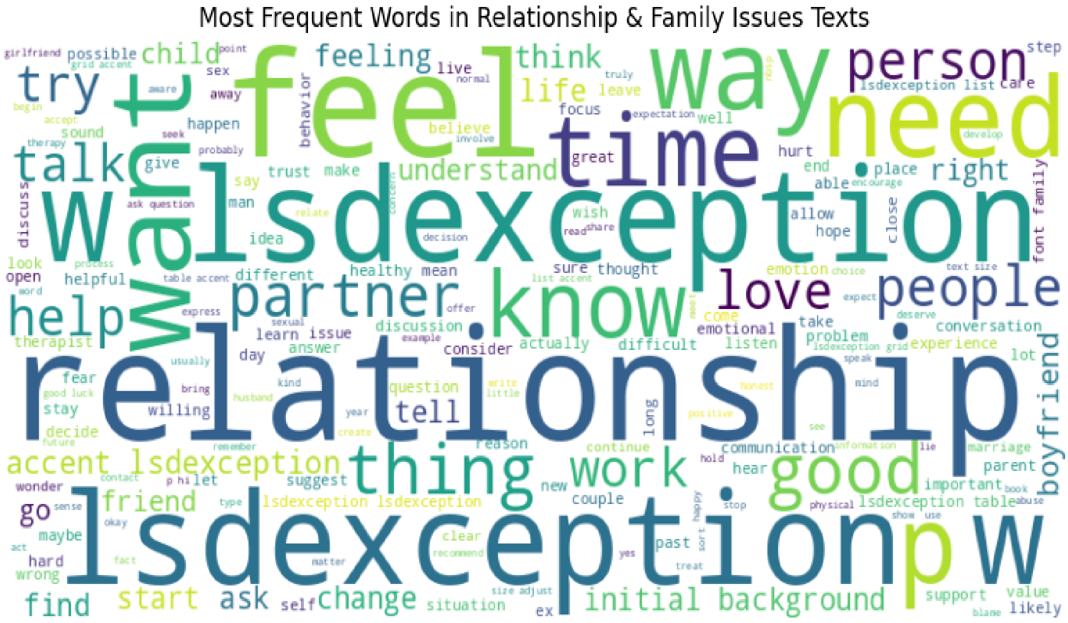


fig 6.8 word cloud for relationship & family issues texts

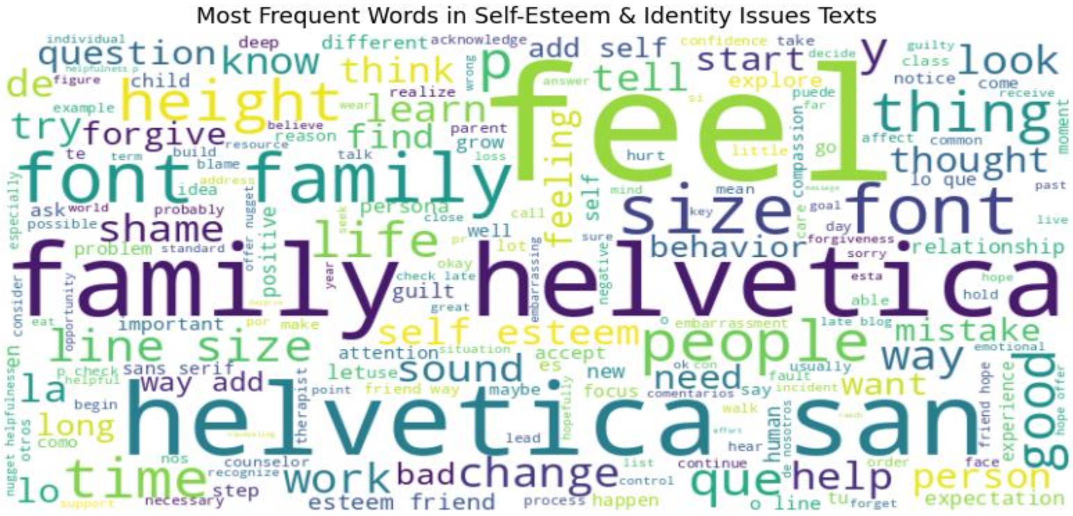


fig 6.9 word cloud for self – Esteem & identity issues texts

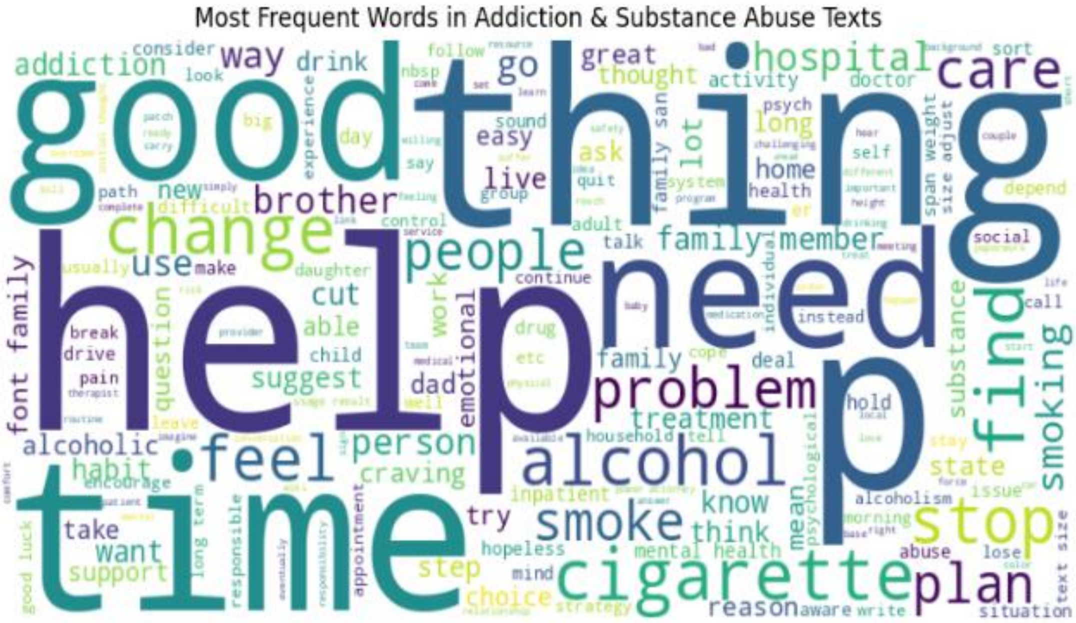


fig 6.10 word cloud for addiction & substance abuse texts

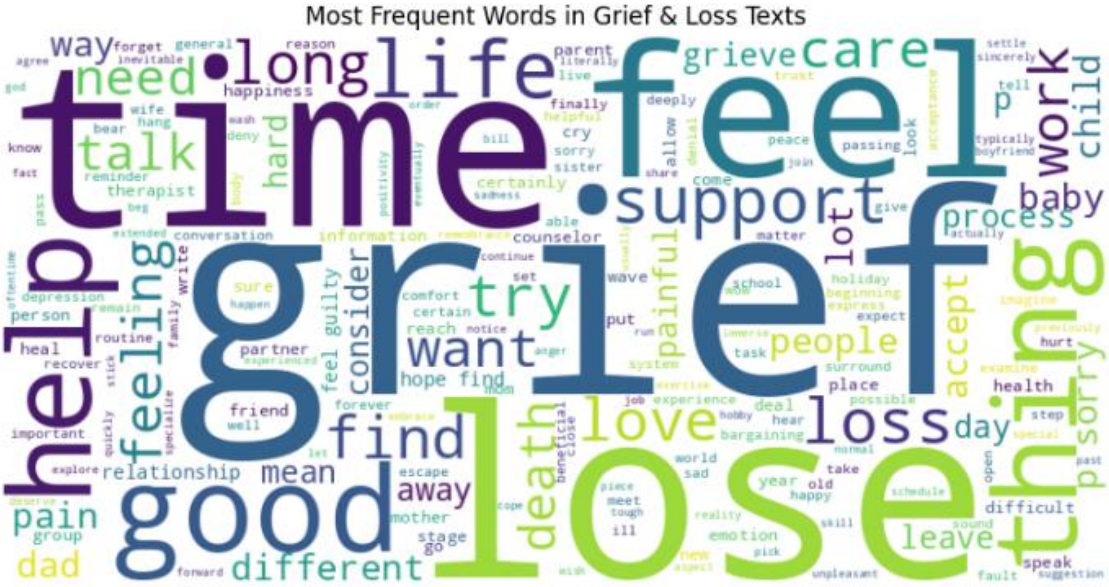


fig 6.11 word cloud for grief & loss texts

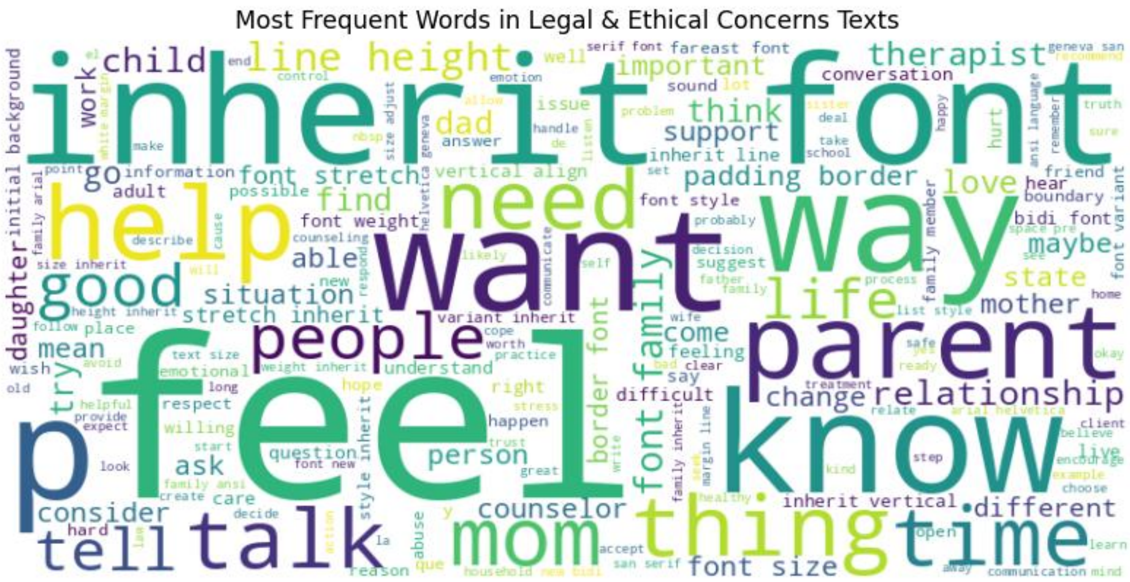
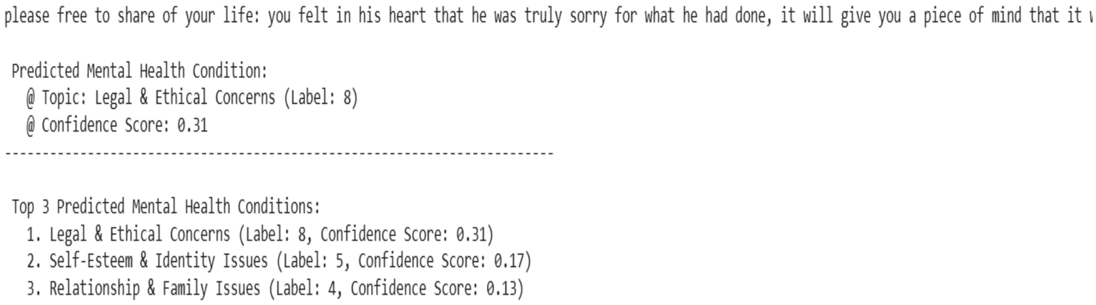


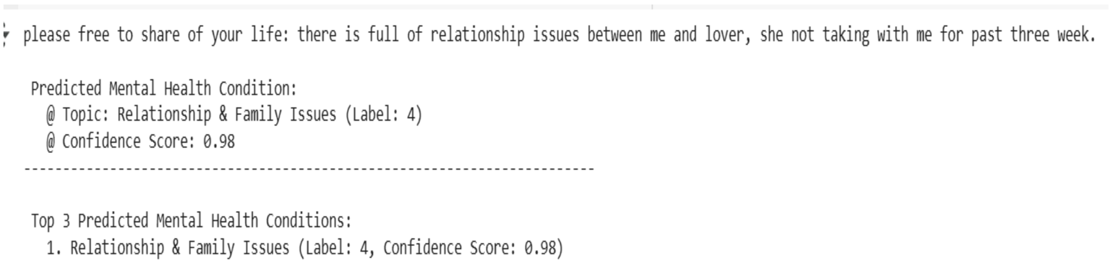
fig 6.12 word cloud for legal & ethical concerns

6.2.4 RESULT



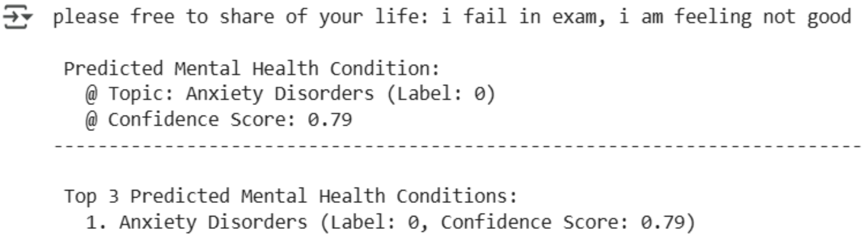
```
please free to share of your life: you felt in his heart that he was truly sorry for what he had done, it will give you a piece of mind that it

 Predicted Mental Health Condition:
   @ Topic: Legal & Ethical Concerns (Label: 8)
   @ Confidence Score: 0.31
-----------------------------------------------------------------------

 Top 3 Predicted Mental Health Conditions:
   1. Legal & Ethical Concerns (Label: 8, Confidence Score: 0.31)
   2. Self-Esteem & Identity Issues (Label: 5, Confidence Score: 0.17)
   3. Relationship & Family Issues (Label: 4, Confidence Score: 0.13)
```

fig 6.13 legal & ethical concerns results



```
please free to share of your life: there is full of relationship issues between me and lover, she not taking with me for past three week.

 Predicted Mental Health Condition:
   @ Topic: Relationship & Family Issues (Label: 4)
   @ Confidence Score: 0.98
-----------------------------------------------------------------------

 Top 3 Predicted Mental Health Conditions:
   1. Relationship & Family Issues (Label: 4, Confidence Score: 0.98)
```

fig 6.14 relationship & family issues results



```
please free to share of your life: i fail in exam, i am feeling not good

 Predicted Mental Health Condition:
   @ Topic: Anxiety Disorders (Label: 0)
   @ Confidence Score: 0.79
--------------------------------------------------------------------------

 Top 3 Predicted Mental Health Conditions:
   1. Anxiety Disorders (Label: 0, Confidence Score: 0.79)
```

fig 6.15 anxiety disorders results

Chapter 7

Results

7.1 RESULT ANALYSIS

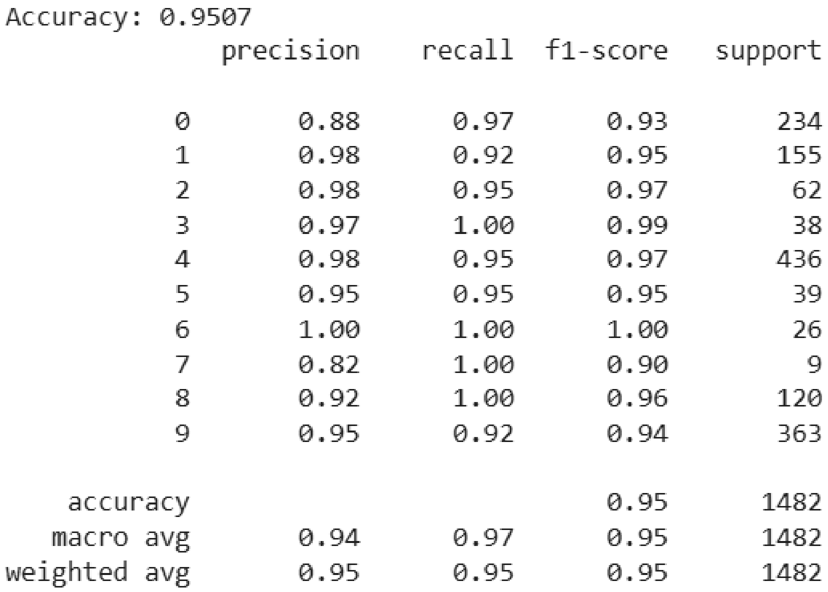
The algorithm correctly predicts mental health disorders using textual input by assessing emotional and contextual signals. In the first case, the input text indicated legal and ethical concerns, which the model detected with a confidence score of 0.31, as well as secondary themes such as self-esteem and family (fig 5.13). In the second scenario, the model identified relationship and family (fig 5.14) concerns with a high confidence score of 0.98, indicating significant emotional discomfort caused by a lack of communication. In the third case, the user reported academic failure and emotional distress, which the model accurately recognized as anxiety disorders (fig 5.15) concerns with a confidence score of 0.79. These findings demonstrate the model's capacity to successfully recognize and categorize numerous mental health conditions via fine-tuned BERT and AdamW optimization.

7.2 CHALLENGES FACED

Several obstacles arose throughout the course of the project, all of which required cautious treatment. One of the most difficult issues was coping with unbalanced data, in which some mental health categories had much less cases than others, possibly impairing model performance and generalization. Another significant challenge was maintaining the semantic clarity of user-supplied content, as many inputs had grammatical mistakes, casual language, or confusing phrasing, making it difficult for models to understand meaning effectively. Furthermore, selecting the appropriate preprocessing techniques and feature extraction methods was crucial for keeping emotional subtleties while minimizing noise.

Training a model capable of distinguishing between closely comparable categories such as "Self-Esteem Issues" and "Relationship Problems" proved difficult owing to emotional context overlap. Finally, refining the model for accuracy and interpretability was an ongoing effort that included testing several architectures and reviewing many metrics to assure dependability in real-world applications.

7.3 EVALUATION METRICS



```
Accuracy: 0.9507
              precision    recall  f1-score   support

           0       0.88      0.97      0.93       234
           1       0.98      0.92      0.95       155
           2       0.98      0.95      0.97        62
           3       0.97      1.00      0.99        38
           4       0.98      0.95      0.97       436
           5       0.95      0.95      0.95        39
           6       1.00      1.00      1.00        26
           7       0.82      1.00      0.90         9
           8       0.92      1.00      0.96       120
           9       0.95      0.92      0.94       363

    accuracy                           0.95      1482
   macro avg       0.94      0.97      0.95      1482
weighted avg       0.95      0.95      0.95      1482
```

fig 7.1 performance matrix & accuracy

The model's performance evaluation shows very encouraging results. With an overall accuracy of 95.07%, the classifier is great at properly recognizing various mental health disorders. Precision, recall, and F1-scores are consistently excellent across virtually all categories, with certain labels, such as Class 6 and Class 3, receiving perfect 1.00 values, indicating flawless prediction in those areas. The macro and weighted averages of F1-score both equal 0.95, demonstrating the model's good generalization across all classes, independent of distribution. Notably, the model performs well with unbalanced data, retaining good recall even in classes with minimal support, such as Class 7 and Class 6. These findings demonstrate the model's resilience and dependability in recognizing a wide range of mental health disorders with high accuracy.

CONCLUSION

This study demonstrates how a fine-tuned BERT model combined with the AdamW optimizer can predict mental health issues using user-provided text inputs. Thanks to powerful natural language processing algorithms, the system can accurately identify a wide range of psychological issues. The experimental results show that the model can detect subtle emotional and environmental patterns, making it a useful tool for early mental health evaluation. With additional refinement and integration with mental health platforms, this approach can assist clinicians in providing timely and individualized care, resulting in improved emotional well-being and early intervention. Furthermore, the model's strong generalization ability across various types of user inputs indicates its suitability for real-world use. Continuous training on diverse and updated datasets allows the model to adapt to changing language patterns and cultural expressions of mental health symptoms. Integration with explainability techniques such as SHAP or LIME improves transparency and builds trust between clinicians and patients. The system is also scalable for use in mobile applications and chat-based therapy services, making it more accessible to underserved or remote populations. Future enhancements, such as emotion trajectory tracking and multimodal analysis that combines text, voice, or facial emotion recognition, could significantly improve its diagnostic capabilities. Overall, this study is a significant step toward creating AI-powered mental health support tools that are proactive, empathetic, and personalized.

LIMITATIONS

The model's capacity to reliably forecast less-represented illnesses may be hampered by class imbalance, which occurs when certain mental health categories have more data than others.

BERT is a black-box model with good accuracy but limited interpretability. Without integrated explainability tools, doctors may struggle to comprehend why a certain prognosis was made.

The model was trained on a limited sample of treatment transcripts, which may not adequately reflect diverse mental health manifestations across cultures, age groups, or socioeconomic backgrounds. This can reduce the model's generalizability.

The model processes text in fixed-length segments (typically 512 tokens), potentially losing important context during longer conversations.

Handling sensitive mental health data raises concerns about privacy, consent, and ethical considerations. Data security and anonymity are critical for real-world deployment.

FUTURE WORKS

Multilingual accommodate: Extend the approach to accommodate several languages, allowing for greater accessibility and inclusion among varied communities.

Explainability and Interpretability: Use explainable AI (XAI) tools like SHAP or LIME to enable users and mental health experts comprehend the foundation of model predictions.

Hybrid Modeling Approaches: Combine BERT-based categorization with psychological knowledge graphs or rule-based systems to enhance the model's contextual awareness.

Clinical Validation: Work with mental health practitioners to evaluate predictions against clinical diagnoses and increase model dependability in real-world contexts.

Privacy-Preserving Techniques: Use techniques such as federated learning or differential privacy to secure user data while preserving model performance.

Real-Time Chatbot Integration: Develop a real-time conversational agent that uses this model for live mental health support and immediate feedback.

REFERENCE

[1]. Sridhar, S., Kazi, N., Kahanda, I., & McCrory, B. Psychiatry transcript annotation: Process study and improvements. In *Proceedings of the International Symposium on Human Factors and Ergonomics in Health Care* 2021, June Vol. 10, No. 1, pp. 71-75.

[2]. Hawa, S., Akella, S., Kaushik, S., Joshi, V., & Kalbande, D. Analysis of therapy transcripts using natural language processing. *International Journal of Engineering and Advanced Technology (IJEAT),*2020, vol *9*(6), pp. 489–494.

[3]. Chung, J., & Teo, J. Mental health prediction using machine learning: taxonomy, applications, and challenges. *Applied Computational Intelligence and Soft Computing*, *2022*(1), pp. 1-19.

[4]. Yu, J. S., Xue, A. Y., Redei, E. E., & Bagheri, N.A support vector machine model provides an accurate transcript-level-based diagnostic for major depressive disorder. *Translational psychiatry,* 2016, vol *6*(10), pp. 931-931.

[5]. Hahn, T., Nierenberg, A. A., & Whitfield-Gabrieli, S. Predictive analytics in mental health: applications, guidelines, challenges and perspectives. *Molecular psychiatry*,2017, vol *22*(1), pp. 37-43.

[6]. Kazi, N. H. *Automated clinical transcription for behavioral health clinicians* (Doctoral dissertation, Montana State University-Bozeman, College of Engineering),2022, pp.1-90.

[7]. Kaynak, E. B., & Dibeklioğlu, H. Systematic analysis of speech transcription modeling for reliable assessment of depression severity. *Sakarya University Journal of Computer and Information Sciences*,2024, vol *7*(1), pp. 77-91.

[8]. Malgaroli, M., Hull, T. D., Zech, J. M., & Althoff, T. Natural language processing for mental health interventions: a systematic review and research framework. *Translational Psychiatry*, 2023, vol *13*(1), pp. 309.

[9]. Wu, Y., Chen, J., Mao, K., & Zhang, Y. Automatic post-traumatic stress disorder diagnosis via clinical transcripts: a novel text augmentation with large language models. In *2023 IEEE Biomedical Circuits and Systems Conference (BioCAS)* 2023, October, pp. 1-5.

[10]. Benjet, C., Bromet, E., Karam, E. G., Kessler, R. C., McLaughlin, K. A., Ruscio, A. M., ... & Koenen, K. C.The epidemiology of traumatic event exposure worldwide: results from the World Mental Health Survey Consortium. *Psychological medicine*, 2016,vol *46*(2), pp.327-343.

[11]. Tlachac, M. L., Toto, E., Lovering, J., Kayastha, R., Taurich, N., & Rundensteiner, E. Emu: Early mental health uncovering framework and dataset. In *2021 20th IEEE international conference on machine learning and applications (ICMLA)*, 2021, December, pp. 1311-1318

[12]. Laijawala, V., Aachaliya, A., Jatta, H., & Pinjarkar, V. Mental health prediction using data mining: a systematic review. In *Proceedings of the 3rd international conference on advances in science & technology,2020, April*, pp.157-162

[13]. Squires, M., Tao, X., Elangovan, S., Gururajan, R., Zhou, X., Acharya, U. R., & Li, Y. Deep learning and machine learning in psychiatry: a survey of current progress in depression detection, diagnosis and treatment. *Brain Informatics*, 2023,vol *10*(1), pp. 1-10.

[14]. Nash, C., Nair, R., & Naqvi, S. M. Machine learning in ADHD and depression mental health diagnosis: a survey. 2024, vol *11*. pp. 86297–86317

[15]. Joshi, M. L., & Kanoongo, N. Depression detection using emotional artificial intelligence and machine learning: A closer review. *Materials Today: Proceedings*, 2022, *58*, pp. 217-226.

[16]. Liu, S., Lu, T., Zhao, Q., Fu, B., Wang, H., Li, G., ... & Lyu, N. A machine learning model for predicting patients with major depressive disorder: A study based on transcriptomic data. *Frontiers in Neuroscience*, 2022, vol 16, pp. 1-16.

[17]. Cohen, J., Wright-Berryman, J., Rohlfs, L., Wright, D., Campbell, M., Gingrich, D., ... & Pestian, J. A feasibility study using a machine learning suicide risk prediction model based on open-ended interview language in adolescent therapy sessions. *International journal of environmental research and public health*, 2022, vol *17*(21), pp. 1-12.

[18]. Qin, J., Liu, C., Tang, T., Liu, D., Wang, M., Huang, Q., & Zhang, R. Mental-Perceiver: Audio-Textual Multi-Modal Learning for Estimating Mental Disorders. In *Proceedings of the AAAI Conference on Artificial Intelligence* 2025, April, Vol. 39, No. 23, pp. 25029-25037.

[19]. Perivolaris, A., Rueda, A., Parkington, K., Soni, A., Rambhatla, S., Samavi, R., ... & Bhat, V. Opinion: Mental health research: to augment or not to augment. *Frontiers in Psychiatry*, 2025, pp. *6*.

[20]. Rasool, A., Aslam, S., Hussain, N., Imtiaz, S., & Riaz, W. nBERT: Harnessing NLP for Emotion Recognition in Psychotherapy to Transform Mental Health Care. *Information*, 2025, vol *16*(4), pp.1–18

[21]. Gutierrez, E. D., Cecchi, G. A., Corcoran, C., & Corlett, P. Using automated metaphor identification to aid in detection and prediction of first-episode schizophrenia.

In *Proceedings of the 2017 conference on empirical methods in natural language processing* 2017, September, pp. 2923-2930.

[22]. Althoff, T., Clark, K., & Leskovec, J.Large-scale analysis of counseling conversations: An application of natural language processing to mental health. *Transactions of the Association for Computational Linguistics*, 2016, *4*, pp. 463-476.

[23]. Alhusen, J. L., Gross, D., Hayat, M. J., Rose, L., & Sharps, P. The role of mental health on maternal-fetal attachment in low-income women. *Journal of Obstetric, Gynecologic & Neonatal Nursing*, 2012, *41*(6), pp. 71-81.

[24]. Lu, J., Liu, B., Lian, Z., Cai, C., Tao, J., & Zhao, Z. (2022, December). Prediction of depression severity based on transformer encoder and CNN model. In *2022 13th international symposium on chinese spoken language processing (ISCSLP)* 2022, December, pp. 339-343.