TAKE YOUR CARE

A PROJECT REPORT
for
Project (KCA451)
Session (2024-25)

Submitted by

MEENAKSHI BHARADWAJ
(2300290140102)
RITAKSHI SINGH
(2300290140141)

Submitted in partial fulfilment of the
Requirements for the Degree of

MASTER OF COMPUTER APPLICATION

Under the Supervision of
Mr. Arpit Dogra
Assistant Professor

Submitted to

DEPARTMENT OF COMPUTER APPLICATIONS
KIET Group of Institutions, Ghaziabad
Uttar Pradesh-201206

(MAY 2025)

CERTIFICATE

Certified that **Meenakshi Bharadwaj (2300290140102)** and **Ritakshi Singh (2300290140141)** have carried out the project work having "**Take Your Care**" (**Project-KCA451**) for **Master of Computer Application** from Dr. A.P.J. Abdul Kalam Technical University (AKTU) (formerly UPTU), Lucknow under my supervision. The project report embodies original work, and studies are carried out by the students themselves and the contents of the project report do not form the basis for the award of any other degree to the candidate or to anybody else from this or any other University/Institution.

This is to certify that the above statement made by the candidate is correct to the best of my knowledge.

Date:

Mr. Arpit Dogra	**Dr. Akash Rajak**
Assistant Professor	**Dean**
Department of Computer Applications	**Department of Computer Applications**
KIET Group of Institutions, Ghaziabad	**KIET Group of Institutions, Ghaziabad**

Take Your Care

ABSTRACT

In modern society, rapid technological advancements and increased digital connectivity have reshaped human interactions, work dynamics, and lifestyle patterns. While these advancements offer numerous benefits, they have also contributed to the rise of mental health issues such as anxiety, depression, and stress-related disorders. The growing prevalence of mental health conditions, coupled with social stigma and limited access to professional care, has created a need for innovative, technology-driven solutions.

In recent years, individuals have increasingly turned to digital platforms such as social media, blogs, and online forums to express their thoughts and emotions. This presents an opportunity to leverage advances in Natural Language Processing (NLP) to detect signs of mental health conditions through text analysis.

This research explores the development of an AI-powered text-based emotion analysis model designed to identify potential mental health disorders. By leveraging Natural Language Processing (NLP) and machine learning techniques, the model analyses user-generated text to detect conditions associated with mental issues like depression, anxiety. The system integrates transformer-based architectures such as BERT, RoBERTa, trained on publicly available mental health datasets, to achieve high accuracy in classification.

The proposed solution aims to bridge the gap between individuals experiencing mental health concerns and professional intervention by providing preliminary insights and recommendations. While the model does not replace clinical diagnosis, it serves as a supportive tool for early detection and awareness. This research contributes to the intersection of AI and mental health, emphasizing the importance of ethical AI applications in addressing modern psychological challenges.

ACKNOWLEDGEMENTS

Success in life is never attained single-handedly. My deepest gratitude goes to my project supervisor, **Mr. Arpit Dogra (Assistant Professor)** for his guidance, help, and encouragement throughout my project work. Their enlightening ideas, comments, and suggestions.

Words are not enough to express my gratitude to **Dr. Akash Rajak**, **Professor and Dean**, **Department of Computer Applications**, for his insightful comments and administrative help on various occasions.

Fortunately, I have many understanding friends, who have helped me a lot on many critical conditions.

Finally, my sincere thanks go to my family members and all those who have directly and indirectly provided me with moral support and other kind of help. Without their support, completion of this work would not have been possible in time. They keep my life filled with enjoyment and happiness.

Meenakshi Bharadwaj

Ritakshi Singh

…………

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

Abbreviation	Definition
AI	Artificial Intelligence
BERT	Bidirectional Encoder Representations from Transformer
GPU	Graphical Processing Unit
ML	Machine Learning
NLP	Natural Language Processing
RoBERTa	Robustly Optimized BERT Approach
TPU	Tensor Processing Unit

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

1.1 BACKGROUND

Mental health disorders are becoming an increasingly pressing issue around the globe, impacting millions of people from all walks of life and age groups. As stress, anxiety, and depression rise—particularly among students and working professionals—it's more important than ever to catch these issues early and intervene. Traditionally, mental health assessments have depended on self-reported surveys and clinical interviews, which can take a lot of time and may not always be reliable.

However, recent breakthroughs in Natural Language Processing (NLP) and machine learning are paving the way for automated mental health detection. By analyzing social media posts, online forums, and other digital interactions, we can tap into a wealth of textual data that reveals signs of mental health challenges.

Transformer models such as BERT, RoBERTa (Robustly Optimized BERT Approach) model with their context-aware capability have demonstrated outstanding capabilities in grasping contextual language, making it an excellent tool for identifying mental health conditions through text analysis. These models are cutting-edge NLP transformer models with the following advantages:

- Performance
- Reliability
- Scalability

1.2 MOTIVATION

In today's fast-paced and digitally connected world, mental health challenges are becoming increasingly prevalent. Millions silently suffer from conditions like depression, anxiety, bipolar disorder, and other psychological struggles—often without access to timely diagnosis or professional support. The stigma surrounding mental health, coupled with limited access to resources, creates barriers that prevent individuals from seeking help.

This project is driven by a desire to harness the power of artificial intelligence to bridge that gap. By analysing the way individuals express their emotions through text, we aim to develop a system that can understand and identify mental health issues early and accurately. Transformer-based language models, especially those fine-tuned for emotional or mental health content, offer a powerful opportunity to interpret subtle cues in language that traditional methods might miss.

But the goal doesn't stop at detection. We envision a system that not only recognizes when someone is struggling but also guides them toward appropriate help—connecting users with mental health professionals in a timely, personalized manner. By combining the strength of cutting-edge NLP with compassionate application design, this project aspires to be more than just a technical achievement—it aims to make a real-world impact, one message at a time.

Ultimately, this research is motivated by empathy, innovation, and the belief that technology can—and should—play a role in building a healthier, more supportive society.

1.3 OBJECTIVES

The primary goal of this research is to develop a machine learning model that is capable of detecting mental health conditions based on user-submitted text. This model aims to tackle significant challenges in evaluating mental health by utilizing the latest advancements in natural language processing (NLP) techniques. The specific objectives are as follows:

1. **Automating Mental Health Screening**

   When it comes to mental health assessments, we often depend on self-reported surveys and face-to-face evaluations. While these methods are valuable, they can be pretty time-consuming and resource-heavy, not to mention they can be influenced by how people choose to respond. This project is all about streamlining that process. By diving into textual data—thoughts and feelings of a person elaborated through a text—we aim to spot signs of depression, anxiety, and stress. With the help of AI, we can offer quick, scalable, and budget-friendly preliminary mental health evaluations, which could really lighten the load on our healthcare systems.

2. **Early Detection and Intervention**

   Spotting mental health issues early on is key to getting the right treatment. This model is designed to sift through language patterns, emotional tones, and meaning in text to catch potential mental health problems before they get worse. For instance, signs like feelings of hopelessness, pulling away from social interactions, or a persistent negative outlook might suggest depression. By picking up on these early warning signs, the model can prompt individuals to seek professional help sooner, which can lead to better long-term results for those who are at risk.

3. **Improving Accessibility and Reducing Barriers to Care**

   Many people encounter obstacles when trying to access mental health services. These challenges often stem from factors like cost, stigma, or simply a lack of available resources. This innovative AI-driven tool can seamlessly integrate into mobile apps, telehealth platforms, and educational institutions, offering confidential and instant mental health insights. By enhancing the accessibility of mental health screenings, this project strives to close the gaps in mental healthcare access.

   This initiative is all about using AI to change the way we identify mental health problems. Our goal is to speed up, improve accuracy, and make the procedure more widely available. By giving professionals and individuals data-driven tools that facilitate early intervention and support, we hope to reduce the burden on the world's mental health system.

1.4 KEY FEATURES

The primary features of the project are as follows:

1. **RoBERTa-based Text Classification**

   The model uses the powerful RoBERTa (Robustly Optimized BERT Approach) model, which has been fine-tuned for mental health detection. Unlike standard NLP models, RoBERTa is highly effective in contextual language understanding, picking up on subtle emotional hints, and decoding intricate sentence structures. It processes user-generated content to categorize mental states as depression, anxiety, stress, or neutral with high accuracy.

   The transformer-based model architecture of our model allows it to handle sarcasm, metaphors, and implicit expressions of distress commonly seen in discussions on mental health. This advanced categorization is the basis of our detection system.

2. **Multi-Label Mental Health Detection**

   It understands that people often deal with overlapping challenges, like battling depression alongside anxiety or managing stress while facing sleep disorders. Instead of just telling you whether a condition is present or not, the model gives probability scores for each identified issue, shedding light on how severe they are. This approach allows for a more thorough evaluation that truly captures the complexities of mental health in the real world. Clinicians get detailed insights that illustrate how various conditions might be interacting in each unique case.

3. **Sentiment and Emotion Analysis**

   Alongside our diagnostic classification, our system dives deep into sentiment and emotion analysis of the text you provide. It goes beyond simply identifying whether the language is positive, negative, or neutral; it also pinpoints specific emotions such as sadness, anger, fear, or hopelessness. This two-layered approach offers a richer understanding of the user's mental state, moving past just clinical labels. For example, it can help differentiate between agitated depression and melancholic depression. These insights are incredibly useful for both automated assessments and professional evaluations.

4. **User-Friendly Interface**

   Our system features a user-friendly interface that allows individuals to share their feelings through straightforward text input. It quickly analyzes those emotions and provides a clear, easy-to-understand mental health assessment. The results come with simple, non-technical feedback and tailored wellness suggestions. The design focuses on creating a supportive, stigma-free environment, featuring soothing colors and easy navigation. This approach ensures that monitoring mental health is both accessible and engaging for everyone.

1.5 SCOPE OF THE PROJECT

- **Target User Base**
  - The system is designed to serve individuals seeking mental health self-assessment, including students, working professionals, and vulnerable populations.
  - Mental health practitioners can use it as a screening tool to identify at-risk patients during preliminary consultations.
  - Organizations like universities and corporations may integrate it into their wellness programs for employee/student support.

  The platform particularly targets high-stress groups including healthcare workers, first responders, and remote employees who face elevated mental health risks.

- **Functional Scope**
  - The project focuses on analyzing user-generated text to detect signs of common mental health conditions through natural language processing.
  - It provides multi-label classification for conditions like depression, anxiety and stress while distinguishing neutral states.
  - The system delivers real-time analysis with visual feedback about emotional states and risk levels. Importantly, it serves as a screening aid rather than a diagnostic tool, always recommending professional verification for concerning results.

- **Technical Scope**
  - Development involves fine-tuning the RoBERTa model using curated mental health datasets from Reddit, Twitter and clinical sources.
  - The system incorporates sentiment analysis layers to detect emotional tones and linguistic patterns associated with psychological distress.
  - For deployment, we're building a secure web interface with potential for mobile integration, using encrypted cloud processing to maintain privacy.

  The architecture allows for future expansion including multilingual support and integration with health record systems.

- **Future Expansion Scope**

There are exciting possibilities for expanding our capabilities. We could integrate multimodal analysis that takes into account voice tone and facial expressions through various APIs. Imagine a system that evolves into a predictive tool, tracking mental health trends over time. There's also a great opportunity to connect with IoT devices and wellness apps, offering a comprehensive approach to mental health monitoring. Future iterations might even include personalized coping strategies powered by generative AI. Plus, with a modular design, we can keep enhancing the system as NLP technology and mental health research continue to progress.

CHAPTER 2

PROBLEM IDENTIFICATION & FEASIBILITY STUDY

2.1 PROBLEM IDENTIFICATION

Mental health disorders affect a significant portion of the global population, yet many cases remain undetected due to lack of awareness, stigma, or limited access to healthcare professionals. Traditional diagnosis often relies on clinical interviews, which can be time-consuming, subjective, and inaccessible for many individuals.

With the increasing digital expression of thoughts and emotions through social media, forums, and personal texts, there exists an opportunity to analyse language patterns and identify early signs of mental health issues using Natural Language Processing (NLP). However, the problem lies in building an accurate, robust, and scalable system that can —

- Understand subtle emotional expressions in text.
- Differentiate between various mental health conditions.
- Provide real-time feedback and support recommendations.

2.1.1 Growing Mental Health Crisis

- The world is facing a growing mental health crisis, with rising rates of depression, anxiety, and stress disorders, particularly among young people and working professionals.
- Traditional methods for diagnosing these issues, like clinical interviews and questionnaires, can be really time-consuming, expensive, and often hard to access.

2.1.2 Barriers to Early Detection

- A lot of people hesitate to seek help because of stigma, a lack of awareness, or simply not having access to mental health services.
- Many subtle early symptoms can easily be overlooked until they escalate into more serious conditions.

2.1.3 Limitations of Current Digital Solutions

- Most mental health apps out there depend heavily on users manually inputting their data, like tracking their moods or filling out surveys, instead of using passive analysis.

- Unfortunately, many of these apps don't have the clinical-grade accuracy needed to effectively detect mental states through natural language.

2.1.4 Need for Proactive Intervention

- When diagnoses are delayed, it often leads to worse outcomes and higher treatment costs.

- Currently, there isn't a widely accepted AI tool that can provide automated, real-time mental health screenings based on everyday communication.

2.2 FEASIBILITY STUDY

A feasibility study is an essential part of any project as it helps assess whether the project is viable and whether it can be successfully executed within the given constraints. For "Take Your Care" Application, the feasibility study is divided into four key categories:

- Economic Feasibility
- Technical Feasibility
- Operational Feasibility
- Behavioural Feasibility

2.2.1 Economic Feasibility

The economic feasibility of this project refers to the cost-effectiveness of implementing the mental health detection system using transformer-based models like BERT, RoBERTa.

- **Cost-effective Development**: Most tools used are open-source. Google Colab provides free access to GPUs for model training.

- **Scalability**: Once trained, models can be deployed on cloud platforms with minimal maintenance costs.

2.2.2 Technical Feasibility

Technical feasibility assesses whether the required technologies, expertise, and infrastructure are available to implement the project successfully.

- **Availability of Tools**: Pretrained transformer models (BERT, RoBERTa) are openly available via Hugging Face and can be fine-tuned on custom datasets.

- **Infrastructure**: Model training and evaluation are feasible using Google Colab, which provides GPU/TPU acceleration.

- **Libraries**: Python libraries such as PyTorch, TensorFlow, Hugging Face Transformers, and scikit-learn facilitate implementation.

- **Model Availability:** State-of-the-art transformer models like BERT, RoBERTa are readily accessible via Hugging Face with pre-trained weights and easy-to-use APIs.

- **Hardware Support:** Google Colab provides sufficient GPU/TPU resources for model training, embedding generation, and fine-tuning on datasets of up to 10000 samples.

- **Extensibility:** The system allows easy incorporation of additional models (like DistilBERT or ClinicalBERT) or features such as explainability modules or therapist recommendations.

2.2.3 Operational Feasibility

Operational feasibility evaluates whether the system will function effectively in the real world and meet user expectations.

- **Functionality:** The system is capable of predicting the mental health status from user-submitted text and also recommends therapist based on city preference when status is not normal.

- **Usability:** With the integration of symptom explanation and therapist suggestion modules, the system becomes user-friendly for both clinicians and individuals.

- **Maintenance:** Routine updates to the models or datasets can be handled easily without major architectural changes.

- **Use Case Integration:** The system can be integrated into a web application and with advancements, it can be integrated into mobile application as well, where users input text and receive mental health predictions and therapist recommendations.

- **User Experience:** Non-invasive, text-based input ensures ease of use and lowers the barrier for mental health support access.

The proposed system is operationally feasible and can be smoothly adopted in academic, research, or preliminary clinical environments.

2.2.4 Behavioral Feasibility

Behavioral feasibility examines how users (clinicians, patients, developers) will respond to the new system.

- **User Trust and Privacy**: By avoiding sensitive personal identifiers and using anonymized text inputs, the system can ensure privacy compliance, encouraging user trust.

- **Adaptability:** Therapists and mental health professionals can adapt to the system with minimal training

- **Perceived Usefulness**: Since the tool offers a faster, scalable way to triage mental health concerns, it is likely to be viewed as beneficial rather than threatening by professionals.

With user-centric design and privacy-focused architecture, behavioral feasibility is high, and adoption likelihood is favorable.

2.3 CONCLUSION

Based on the analysis of technical, operational, economic, and behavioral aspects, the proposed system for mental health prediction using transformer-based NLP models is highly feasible. The availability of powerful pre-trained models, accessible computational resources, and robust development tools supports the technical implementation. Operationally, the application offers a user-friendly and non-invasive interface for real-time mental health screening and support. Economically, the use of open-source tools and cloud-based resources ensures low development and deployment costs. Moreover, ethical and privacy considerations are thoroughly addressed to ensure responsible usage. Therefore, the project is not only achievable but also scalable and impactful in addressing the growing need for early mental health detection and support.

CHAPTER 3

LITERATURE REVIEW

In recent years, the intersection of artificial intelligence (AI) and mental health research has attracted significant attention, particularly in leveraging Natural Language Processing (NLP) to identify mental health conditions through text analysis. A growing body of literature supports the feasibility and effectiveness of machine learning models in detecting mental health issues from user-generated content, such as social media posts, online forums, and self-reported narratives.

3.1 RELATED WORK

There have been numerous studies in the field of mental health awareness and detection by utilizing the advanced technology and read human brain patterns to help detect the causes of these disorders. The evolution of the study is discussed below:

3.1.1 Traditional Approaches to Mental Health Detection

Earlier efforts in mental health detection primarily utilized rule-based or classical machine learning techniques such as Naive Bayes, Support Vector Machines (SVM), and Random Forests. These models often relied on handcrafted features such as term frequency-inverse document frequency (TF-IDF), sentiment scores, or psychological lexicons like LIWC (Pennebaker et al., 2001). While they demonstrated moderate success, their performance was limited by their inability to fully capture contextual and semantic nuances in language.

3.1.2 Deep Learning and NLP Evolution

The introduction of deep learning models, especially Recurrent Neural Networks (RNNs), LSTMs, and CNNs, brought significant improvements in understanding temporal patterns and syntactic structures. Studies like Shen et al. (2017) used deep LSTM architectures for depression detection, showing that emotional expression can be learned directly from sequences of words. However, these models still struggled with long-range dependencies and context representation.

3.1.3 Transformer Models in Mental Health Applications

The advent of transformer-based architectures revolutionized NLP, with BERT (Devlin et al., 2019) setting a new benchmark for various text classification tasks. BERT introduced bidirectional context understanding, which was shown to be particularly effective in emotion and mental state detection. Studies such as Losada and Crestani (2016) and Yates et al. (2017) applied BERT to suicide risk assessment and depressive language detection with promising results.

RoBERTa (Liu et al., 2019), an optimized version of BERT, further improved model robustness by training on more data and removing certain training constraints. Several studies (e.g., Ji et al., 2022) demonstrated that RoBERTa outperformed BERT in mental health classification tasks due to better generalization and fine-tuned pre-training techniques.

3.1.4 Domain Specific Models: MentalBERT, MentalRoBERTa

Recent research has emphasized the importance of domain-specific language models. **MentalBERT** and **MentalRoBERTa** are transformer models pre-trained on large-scale mental health-related text corpora, such as Reddit threads from mental health subreddits. Ji et al. (2021) introduced MentalBERT and showed that it significantly outperformed general-purpose models on depression, anxiety, and suicidal ideation detection. Similarly, MentalRoBERTa demonstrated superior performance in extracting mental health signals from short and informal text due to its pre-training on context-specific language.

3.2 LIMITATIONS IN CURRENT STUDIES

Despite significant progress, current studies often focus on binary classification (e.g., depressed vs. non-depressed) and rarely explore multi-class mental health classification. Additionally, few works implement a complete end-to-end system that not only detects mental health conditions but also recommends relevant interventions or therapists. Moreover, model comparisons in most studies are limited to one or two models, lacking a broader performance benchmark.

3.3 CONTRIBUTIONS AND IMPROVEMENTS OF THIS STUDY

This study addresses the limitations by:

- Comparing the transformer models (BERT, RoBERTa) on a multi-class mental health classification task (e.g. Depression, Anxiety, Stress, Bipolar Disorder, Normal, etc.)
- Evaluating models using comprehensive metrics including Accuracy, Precision, Recall, etc.
- Deploying the best performing model in a real-world application that detects mental health status and suggest suitable therapists, offering a practical implementation of AI in mental health care.

CHAPTER 4

REQUIREMENT SPECIFICATIONS

The requirement specification outlines the functional and non-functional requirements necessary for the development and deployment of the mental health prediction and therapist recommendation system. The system is based on transformer-based NLP models and aims to classify mental health conditions based on user-provided text inputs.

4.1 FUNCTIONAL REQUIREMENTS

The functional requirements define the core functionalities that a system must provide. The core functionalities of the "**Take Your Care**" web application are listed as follows:

- **User Input Interface**
  - The system must allow users to input free-form text describing their thoughts or emotions.
- **Text Preprocessing Module**
  - The system must clean and preprocess the input text (e.g., removing special characters, converting to lowercase, tokenization).
- **Model Inference Engine**
  - The system must use a trained transformer model (BERT, RoBERTa) to classify the input text into one of the predefined mental health categories (e.g., Normal, Depression, Anxiety, Stress, Bipolar, Suicidal, Personality Disorder).
- **Result Display**
  - The system must display the predicted mental health condition to the user in an understandable format.
- **Therapist Recommendation Module**
  - Based on the predicted category, the system should recommend suitable therapists or suggest next steps (e.g., online counseling platforms or emergency helplines for high-risk cases).
- **Admin Module (Optional)**
  - Admins can update the therapist database or monitor usage logs for insights and improvements.

4.2 NON-FUNCTIONAL REQUIREMENTS

The non-functional requirements define the system's overall qualities, they specify system quality attributes and constraints:

- **Usability**
    - The application must have a user-friendly interface accessible to non-technical users.
- **Performance**
    - The system should return predictions within 1–2 seconds per input on average.
- **Scalability**
    - The backend model should be deployable to cloud infrastructure (e.g., AWS, GCP, or Hugging Face Spaces) to handle multiple users concurrently.
- **Security and Privacy**
    - User inputs must be anonymized and not stored permanently.
    - The system should follow best practices for handling sensitive information.
- **Maintainability**
    - The model and application should be modular and well-documented, allowing updates and retraining as better models or data become available.
- **Portability**
    - The system should be deployable across platforms (web, mobile) with minimal adjustments.

4.3 SYSTEM REQUIREMENTS

The system requirements define the hardware and software specifications essential for the development, training, testing, and deployment of the mental health prediction and therapist recommendation application. This system consists of two main components:

- Phase 1: Model Development and Training Environment
- Phase 2: Deployed Application Environment

4.3.1 Hardware Requirements

The hardware requirements are the system and resources which are to be allocated in order to carry out the project. For the mental detection application project, hardware requirements are discussed for the phases, development and deployment respectively, refer to Tables 4.1 and 4.2 for going through them.

- **Development and Training Phase**

Component	Minimum Requirement
Processor (CPU)	Intel i5
Memory (RAM)	8 GB
Storage	20 GB Free Disk Space
GPU (for training)	GeForce GTX 1650 / T4 (Google Colab)

Table 4.1 Hardware Requirements for Training and Development Phase

- **Deployment Phase**

Component	Minimum Requirement
vCPU	2 or more virtual CPUs
Memory	2 GB RAM (4 GB Recommended)
Disk Space	1-5 GB
Network Bandwidth	Stable internet connection

Table 4.2 Hardware Requirements for Deployment Phase

Recommended Hardware Configuration for Development and Deployment —

- **Processor (CPU)**
  - **Minimum**: Intel Core i5/i7 (10th Gen or newer) or AMD Ryzen 5/7
  - **Recommended**: Intel Core i9 or AMD Ryzen 9 (for quicker preprocessing and multitasking)
  - **Why?**
    - A multi-core CPU with 6 or more cores can really speed things up when it comes to data preprocessing, fine-tuning your models, and running your Flask server.
    - It can handle background tasks like API calls and database queries while your GPU is busy training the model.
- **Graphics Card (GPU)**
  - **Minimum**- NVIDIA RTX 3060 (8GB VRAM) or Google Colab GPU T4
  - **Recommended**- RTX 3080 (10GB+ VRAM) or A100 (40GB VRAM for large-scale training)
  - **Why?**
    - CUDA support is crucial for speeding up deep learning tasks, especially when fine-tuning models like RoBERTa. More VRAM means you can avoid crashes when dealing with large batches of text data. Even a mid-range GPU like the T4 or GTX 1650 can handle real-time inference quite well.
- **RAM**
  - **Minimum**- 16GB DDR4
  - **Recommended**- 32GB DDR4 (or more for massive datasets)
  - **Why?**

- 16GB is enough for moderate-sized datasets, typically around 50k to 100k text samples. If you want to run your Flask server and ML model at the same time or load large embeddings (like RoBERTa-base, which is about 500MB in memory), then 32GB or more is the way to go.

➢ **Storage**

- **Minimum**- 500GB NVMe SSD
- **Recommended**- 1TB NVMe SSD (or consider cloud storage for scalability)
- **Why?**
  - Datasets, especially in mental health research from sources like Reddit and Twitter, can take up 50 to 200GB. A single RoBERTa model might require around 1.5GB, and fine-tuned versions can be even larger. Plus, you'll want fast read/write speeds for logs and databases, which is why SSDs are preferred.

4.3.2 Software Requirements

The software requirements for the development of the application are discussed in Table 4.3 with minimal requirements as well as the requirements used during training and development of the model.

Component	**Requirement**
Operating System	Windows 10/11 or macOS or Linux
Programming Language	Python 3.8+
Development Tools	Jupyter Notebook, Google Colab
Machine Learning Libraries	PyTorch, Hugging Face Transformers, scikit-learn
Transformer Models for training	BERT, RoBERTa
Data Handling Libraries	pandas, numpy
Visualization Libraries	Matplotlib, seaborn
Web Framework (for deployment)	Flask, Ngrok
Text Processing Libraries	nltk, spaCy
Model Deployment	Google Colab (Training), Flask (Deployment)

Table 4.3 Software Requirements for "Take Your Care" App

CHAPTER 5

SYSTEM ANALYSIS

5.1 EXISTING SYSTEM

Traditional methods for detecting mental health issues rely heavily on manual clinical assessments, such as patient interviews and psychological evaluations. While effective, these methods are time-intensive, costly, and inaccessible to many—especially in rural or under-resourced areas. Existing online tools often lack the depth of analysis required to identify multiple types of mental health conditions or provide practical follow-ups such as therapist recommendations.

5.2 PROPOSED SYSTEM

The proposed system aims to:

- Predict mental health conditions (e.g., Depression, Anxiety, Stress, Bipolar, Suicidal) from user-submitted text using transformer-based NLP models.
- Compare multiple state-of-the-art models (BERT, RoBERTa) to determine the best performer.
- Deploy the best-performing model in a web application to classify mental health status.
- Recommend relevant therapists or mental health resources based on the classification outcome.

5.3 SYSTEM ARCHITECTURE

The architecture of the system can be broadly divided into the following layers:

a. Input Layer

- Accepts user-submitted text describing thoughts or emotional states.

b. Preprocessing Layer

- Cleans text (removal of stopwords, punctuation, special characters, etc.).
- Tokenizes and encodes the text using the tokenizer specific to the chosen transformer model.

c. Model Layer

- Performs classification using the selected fine-tuned model.
- Supports dynamic switching between models for evaluation and comparison.

d. Output Layer

- Displays the predicted mental health condition to the user.
- Provides appropriate therapist or support recommendations.

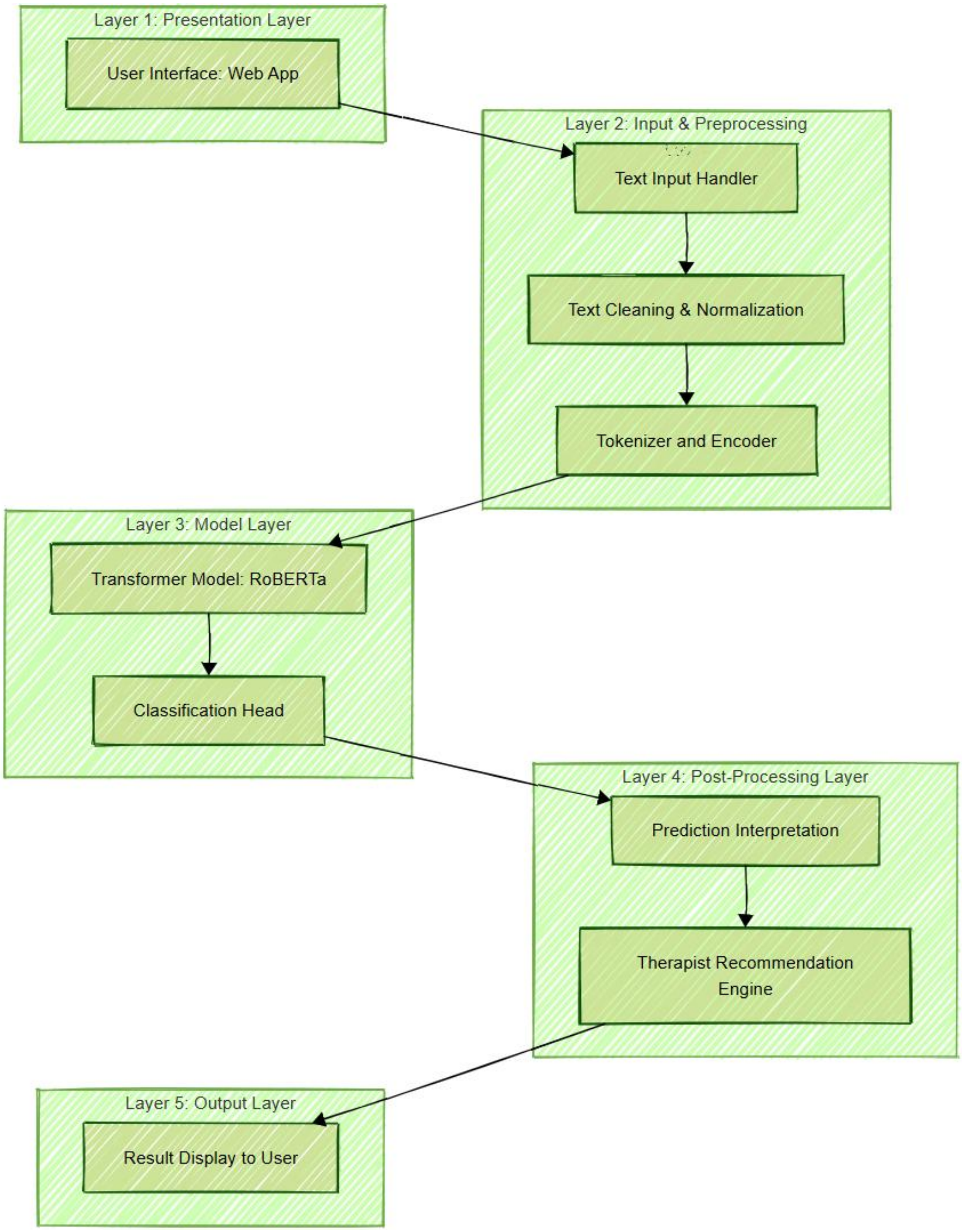

Fig. 5.1 System Architecture

5.4 USE CASE ANALYSIS

All the related use cases are discussed in the Table 5.1 with the actors and their respective use case along with a description of the use case. All possible use cases are consolidated the table. Possible actors interacting with the application are System and User.

The therapists' data is stored in the CSV file. In future, it will be possible to add a new use case for the database management for the therapists' data by the new actor called Admin.

Actor	Use Case	Description
User	Submit mental health description	Inputs text expressing emotional/mental state
System	Predict mental health condition	Classifies input into one of the defined categories
System	Recommend therapist	Suggests appropriate support or resources

Table 5.1 Use Case Table: All Use Cases

5.5 GAP ANALYSIS

The gaps between traditional existing systems and the proposed are thoroughly discussed in the Table 5.2 with all possible varying aspects such as diagnosis approach, classification types, actionable outcomes, accessibility.

Aspect	Existing System	Proposed System
Diagnosis Approach	Manual, subjective	Automated, data-driven NLP model
Classification Types	Often binary or limited	Multi-class classification across 6+ conditions
Actionable Outcomes	None	Therapist recommendation based on prediction
Accessibility	Limited to clinical settings	Available via web interface, scalable access

Table 5.2 Gap Analysis for Existing and Proposed System

5.6 DATA FLOW DIAGRAM

5.6.1 Level-0 DFD (Context Diagram)

Level-0 DFD is a basic diagram that outlines the main flow of the application. It is also known as Context Diagram. Refer to Fig. 5.2 for the context diagram for the proposed application.

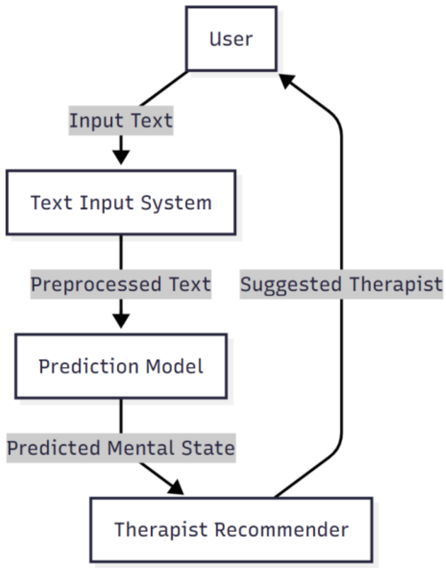

Fig. 5.2 Context Diagram or Level-0 DFD

5.6.2 Level-1 DFD

Level-1 DFD is a diagram that expands the main flow into sub flows of the application. Refer to Fig. 5.3 for the level-1 DFD for the proposed application.

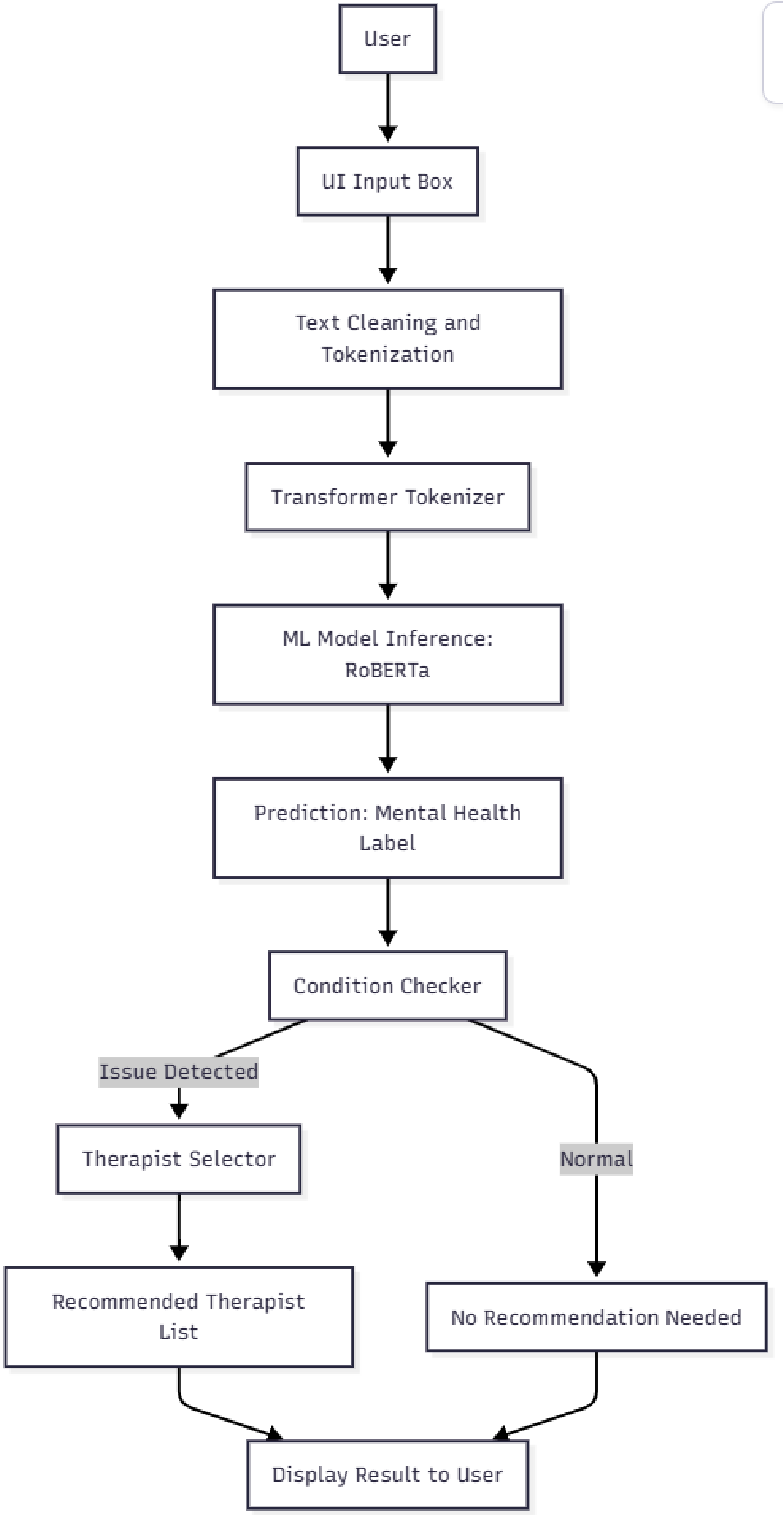

Fig. 5.3 Level-1 DFD

5.7 SEQUENCE DIAGRAM

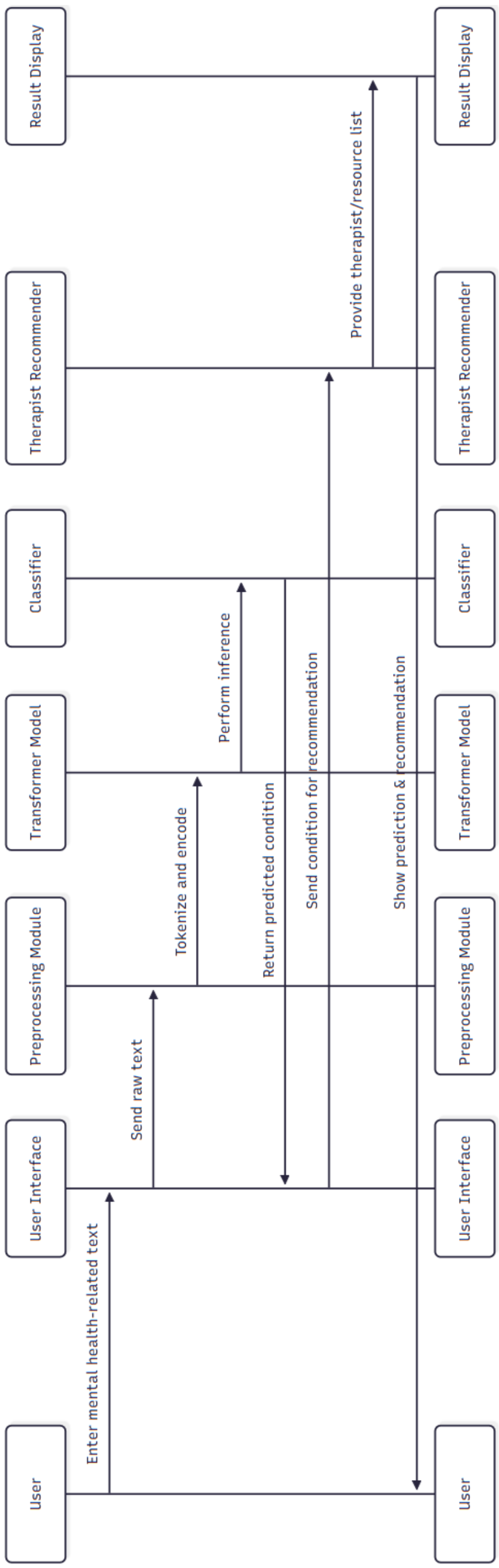

Fig. 5.4 Sequence Diagram

CHAPTER 6

PROJECT PLANNING AND SCHEDULING

Effective project planning is crucial for the timely and organized completion of this mental health detection system. The project was divided into multiple phases, each with defined tasks, deliverables, and timelines. The key stages involved requirement gathering, data preprocessing, model training and evaluation, deployment, and documentation.

6.1 PROJECT PLANNING

A total of seven phases are there for the development of the "**Take Your Care**" project and these phases include requirement analysis, dataset collection and preprocessing, model development, model evaluation, application development, model deployment, documentation and reporting. Proposed time scheduling is given in the Fig. 6.1.

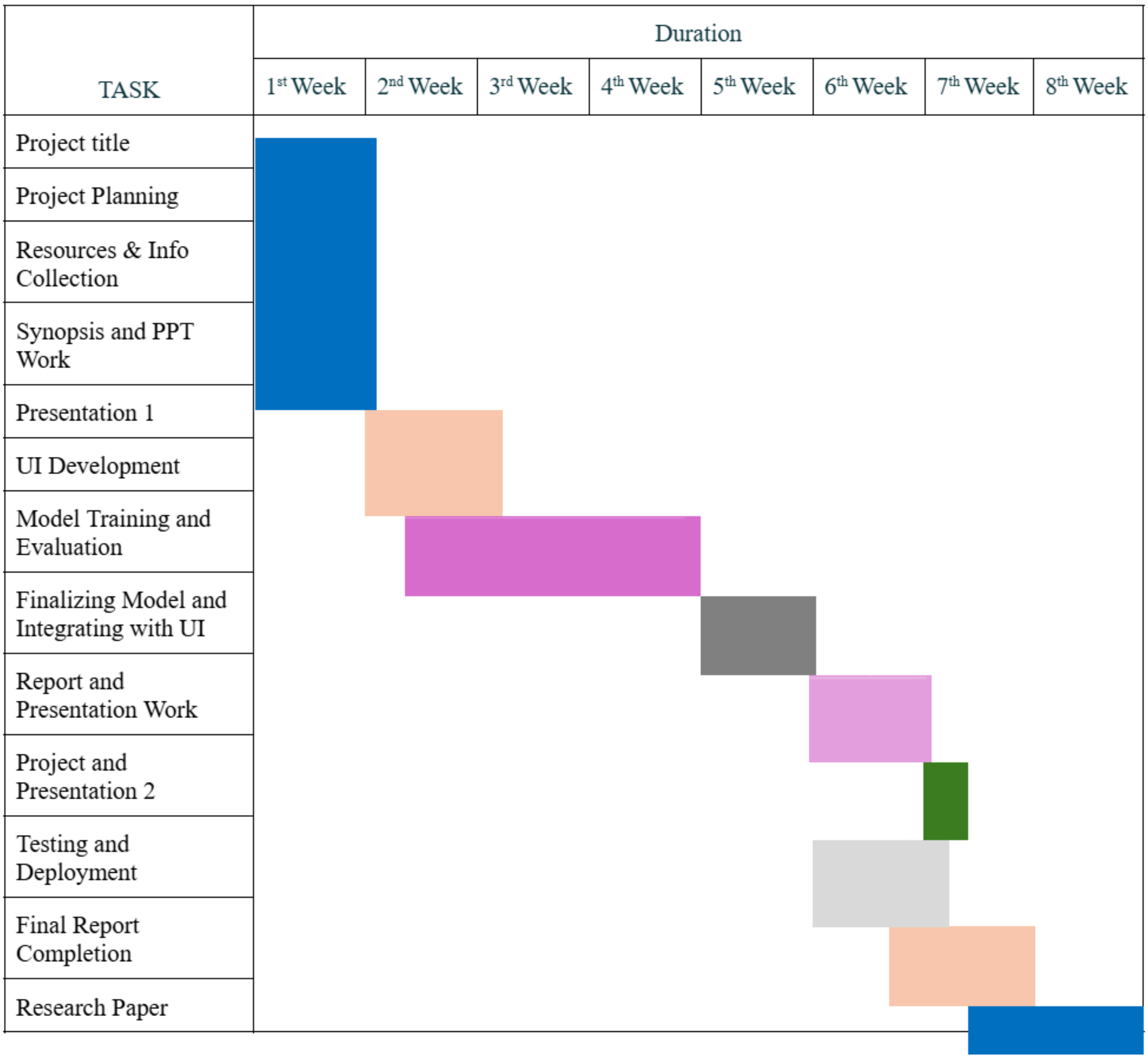

Fig. 6.1 Project Planning before Implementation

6.2 PROJECT SCHEDULING

A Gantt chart for the application “Take Your Care” with proper time allocation is shown in Fig. 6.1 with all the related phases and their completion timelines. It shows the actual timeline of the project during and after the implementation.

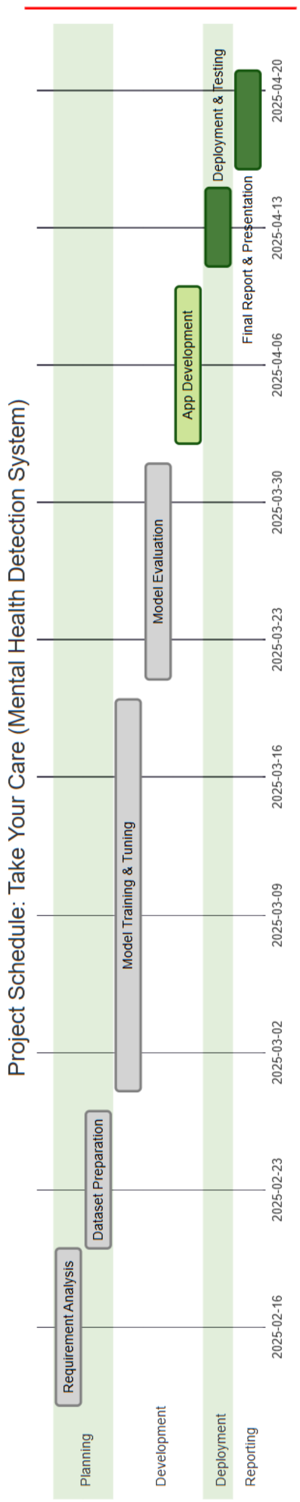

Fig. 6.2 Gantt Chart

CHAPTER 7

NLP AND TRANSFORMER MODELS OVERVIEW

7.1 NATURAL LANGUAGE PROCESSING (NLP)

Natural Language Processing (NLP) is a field of Artificial Intelligence that focuses on the interaction between humans and computers using natural language. It enables machines to understand, interpret, and respond to text or speech in a human-like manner.

In this project, NLP techniques are used to analyse user-submitted text and classify it into predefined mental health categories such as Depression, Anxiety, Suicidal, Bipolar, Stress, etc. Key tasks of NLP involved in this project are:

- Text cleaning and normalization
- Tokenization
- Sequence encoding
- Contextual understanding via Transformers

7.2 TRANSFORMER-BASED MODELS

Transformers are deep learning models introduced by Vaswani et al. in 2017 in the paper *"Attention is All You Need."* They revolutionized NLP by introducing the attention mechanism to understand the context in a sentence more effectively than RNNs or LSTMs.

Core Components of Transformers are:

- **Self-Attention Mechanism**
    - Captures relationships between words in a sentence
- **Positional Encoding**
    - Maintains order information
- **Encoder-Decoder Architecture**
    - Encodes input and decodes output in a parallelized manner.

7.2.1 Models Used

The project used pretrained transformer-based NLP models that includes BERT and RoBERTa. Refer to Fig. 7.1 for the comparative working principle of these models.

- **BERT**

  BERT (Bidirectional Encoder Representations from Transformers) is a powerful natural language processing (NLP) model developed by Google in October 2018.

It's known for its ability to understand context by considering the relationships between words in a sentence, both before and after the target word, making it a state-of-the-art model for various NLP tasks.

- **RoBERTa**

RoBERTa (Robustly Optimized BERT Approach) is based on Google's BERT model. It was released in 2019. It builds on BERT and modifies key hyperparameters, removing the next-sentence pretraining objective.

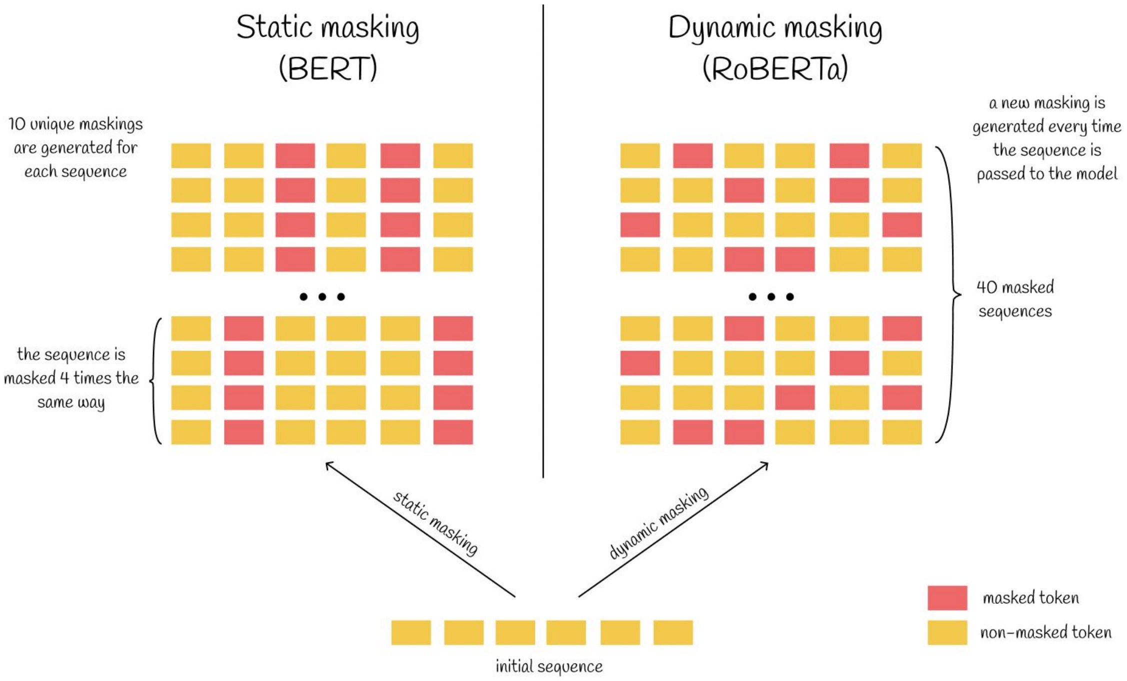

Fig. 7.1 BERT and RoBERTa Transformer Models

7.3 EVALUATION METRICS AND THEIR FORMULAS

To compare the performance of these models objectively, several evaluation metrics were used:

- **Accuracy**

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}$$

  - Measures how often the model's predictions are correct.
  - Suitable for balanced datasets.

- **Precision**

$$\text{Precision} = \frac{TP}{TP + FP}$$

  - The proportion of positive identifications that were actually correct.

- **Recall**

$$\text{Recall} = \frac{TP}{TP + FN}$$

  - The proportion of actual positives that were correctly identified.

- **F1 Score**

$$\text{F1} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

  - Harmonic mean of precision and recall; balances both false positives and false negatives.
  - Balances precision and recall.

CHAPTER 8

PROJECT FLOW

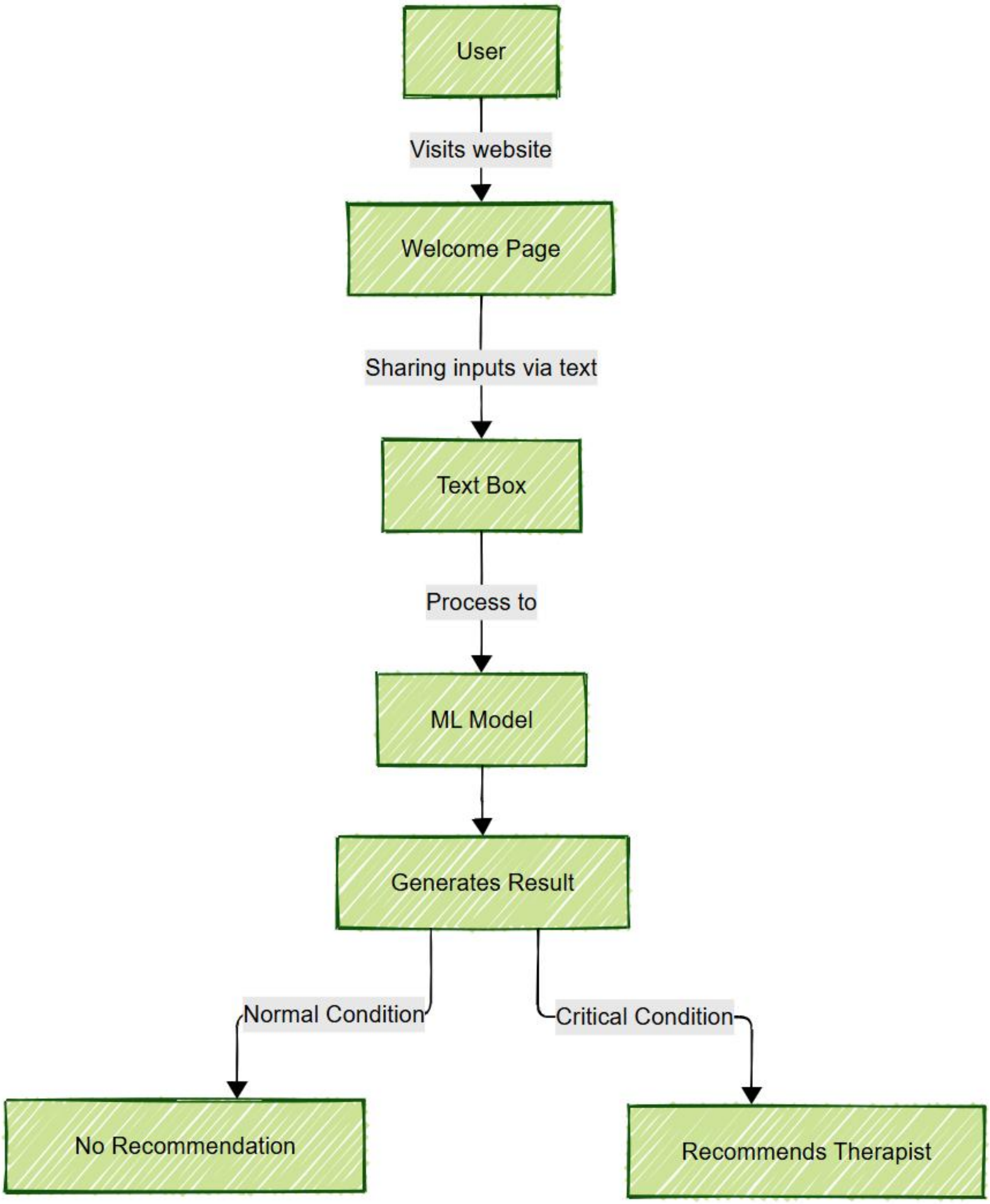

Fig. 8.1 Flow of the Project

The mental health prediction system follows a structured and logical sequence of stages, combining natural language processing and deep learning with user-focused deployment. The flow can be divided into the following phases:

- Data Collection and Preprocessing

- Model Selection and Fine-tuning
- Model Evaluation and Comparison
- Deployment
- Application Integration
    - Therapist Recommendation
    - User Output

8.1 DATA COLLECTION AND PREPROCESSING

8.1.1 Dataset Selection

The dataset used in this project was sourced from a **publicly available mental health dataset on Kaggle**. This dataset includes user-generated text data that reflects various emotional and psychological states, labelled under categories such as:

- Depression
- Anxiety
- Stress
- Suicidal
- Bipolar Disorder
- Personality Disorder
- Normal

8.1.2 Dataset Description

The dataset chosen is "**Sentiment Analysis for Mental Health**" uploaded by **Suchintika Sarkar** on the Kaggle platform for public use. It consists of up to 51K rows with each consisting of three columns:

- **unique_id:** This column captures the unique identifiers for each entry in the dataset.
- **statement:** This column contains free-text collected from the users via online sources. It is the primary column for our project.
- **status:** This is the target column for our project which showcases the actual output for the emotion or sentiment embedded in the statement.

It is available in the CSV format and as per the ethical consideration it is used strictly for the academic and research purposes to promote mental health support technologies.

8.1.3 Dataset Preprocessing

Before training the mental health prediction models, the raw text data underwent several preprocessing steps to improve data quality and ensure compatibility with transformer-based architectures. These steps are essential to clean, normalize, and structure the input for optimal model performance.

1. **Lowercasing:**
    - Text was converted to lowercase to maintain consistency & reduce sparsity.

2. **Removing Punctuation and Special Characters**
    - Unnecessary symbols, URLs, and emojis were removed using regular expressions to eliminate noise.
3. **Stopword Removal**
    - Common words that do not carry significant semantic meaning (e.g., "the", "is", "and") were removed using nltk or spaCy libraries.
4. **Tokenization**
    - Each sentence was split into individual tokens using the tokenizer associated with the transformer model (e.g., BertTokenizer for BERT).
5. **Label Encoding**
    - Mental health categories were encoded into numerical labels for training (e.g., Depression = 0, Anxiety = 1, etc.).
6. **Handling Class Imbalance**
    - Techniques such as:
        - **i.** Over sampling of minority classes
        - **ii.** Under sampling of majority classes
        - **iii.** Weighted loss function during model training
    - For this project **over sampling** and **weighted loss function** techniques were tested wherein the latter approach was finally applied for fair model performance purposes.
7. **Train-Test Split**
    - The dataset was split into training and test sets (e.g., 80:20) using train_test_split() with stratification to preserve class distribution.
8. **Input Formatting for Transformer Models**
    - Each sample was formatted into input IDs, attention masks, and token type IDs as required by Hugging Face models.

The example below shows the text before and after preprocessing. The user-input text before the preprocessing:

Raw Input: *I'm feeling completely hopeless and empty. I don't think I can go on like this.*

After the preprocessing, it processed the text, cleaned it and converted in the form:

Preprocessed Text: ["feeling", "completely", "hopeless", "empty", "think", "go"].

8.2 MODEL SELECTION

The goal of this project is to accurately predict a user's mental health condition based on their textual input. For this task, transformer-based models were chosen due to their state-of-the-art performance in Natural Language Processing (NLP) tasks, especially those involving semantic understanding and contextual language modeling.

8.2.1 Why Transformer-Based Models

- Traditional NLP models (like TF-IDF + SVM or LSTM) struggle with long-range dependencies and contextual comprehension.
- Transformers leverage **self-attention mechanisms**, enabling them to capture relationships between words regardless of their distance in a sentence.
- They are **pre-trained** on large corpora and can be fine-tuned for specific downstream tasks (like mental health classification) with excellent generalization.

8.2.2 Models Selected

The models selected for the project are the NLP transformer models with cutting-edge technologies:

- BERT
- RoBERTa

Model	Base Architecture	Specialization	Why Chosen
BERT	Transformer Encoder	General-purpose language understanding	Benchmark model for comparison
RoBERTa	Optimized BERT	Trained with larger batch size, dynamic masking	Better performance on downstream tasks

Table 8.1 Models Selected with their Features

Table 8.1 discusses about why and what models were selected for the project with their features such as Base Architecture, Specialization and Reason of Selection. After selecting the models, we proceeded with their training, and fine-tuning.

8.3 MODEL TRAINING AND FINETUNING

Fine-tuning is a critical step in adapting a pre-trained transformer model to the specific task of mental health classification. Each of the selected models (BERT, RoBERTa) was fine-tuned using the preprocessed dataset to optimize performance for the target labels.

8.3.1 Dataset Encoding

Transformer models such as BERT, RoBERTa, **do not understand raw human text directly**. Instead, they require numerical inputs to process and learn from. **Dataset encoding** is the essential step that converts raw text into a format that these models can understand.

- Utilized the corresponding tokenizer for each model:
  - **BertTokenizer** for BERT Transformer Model.

    - **RobertaTokenizer** for RoBERTa Transformer Model
- Encoded input text into:
    - input_ids
    - attention_mas
- Used Hugging Face's **datasets.Dataset** and **DataCollatorWithPadding** for efficient batch processing.

8.3.2 Dataset Splitting

The dataset is split by applying **80:20 train-test split** wherein 80% of the data is for the training purpose and 20% of the data is for the testing purpose which helps in evaluating the model based on all metrics such as accuracy, precision, etc.

8.3.3 Training Configuration

As we trained our model using Hugging Face Trainer API, the following configuration was set for the training:

- **Learning Rate**: 2e-5 (adjusted during tuning)
- **Batch Size for Training**: 16
- **Batch Size for Evaluation**: 16
- **Epochs**: 2 to 5 (depends on sample size for training purpose)
- **Evaluation Strategy**: per epoch
- **Save Strategy**: per epoch
- **Load best model at end**: True
- **Logging Steps**: 10

8.3.4 Training Execution

For training execution, we created a trainer object and provided the arguments with the above configuration, model, train_dataset for training, and eval_dataset for evaluation, tokenizer and compute_metrics.

8.3.5 Evaluation Metrics

Post training, each model was evaluated on the test set using the metrices such as **accuracy**, **precision**, **recall**, **F1 score**. Evaluation of metrics is important as it helps to better compare the models and evaluate their performance.

- **Accuracy**: Proportion of correct predictions over total predictions.
- **Precision**: Proportion of correctly predicted positive cases among all predicted positives.
- **Recall**: Proportion of correctly predicted positive cases among all actual positives.
- **F1 Score**: Harmonic mean of precision and recall; especially important for imbalanced datasets.

These metrics together offer a balanced view of each model's classification capability, especially in a sensitive domain like mental health prediction where both false positives and false negatives are significant.

8.4 EVALUATION AND MODEL COMPARISON

This section presents a comprehensive evaluation of the BERT and RoBERTa models based on classification performance metrics such as Accuracy, Precision, Recall, and F1 Score. The models were trained on varying dataset sizes to analyze their learning efficiency and generalization capabilities. A detailed comparison highlights the strengths of each model, guiding the selection of the most suitable one for deployment in a mental health prediction system.

8.4.1 Model Evaluation

This section presents the evaluation of two transformer-based models — **BERT** and **RoBERTa** — used for mental health status classification. The models were assessed based on standard classification metrics: **Accuracy**, **Precision**, **Recall**, and **F1 Score**, across varying training sample sizes (6000, 12000, 24000, and 50000). The purpose was to analyze how both models perform with increasing data and to determine the most effective model for real-world deployment.

8.4.2 Evaluation Results

After evaluating models on different sample sizes, the results were recorded in tabular format. Table 8.2 discusses sample size-wise performance of the BERT model, and the Table 8.3 discusses the sample size-wise performance of the RoBERTa model for all the chosen metrics of evaluation.

- **BERT Performance**

  Refer to Table 8.2 for the BERT model performance at different sample size. It is clearly visible from the performance table that the best observation for the BERT model is noted when the sample size is set to 50000 with its highest evaluation metrics.

Sample Size	**Accuracy**	**Precision**	**Recall**	**F1 Score**
6000	70.33	74.22	70.33	70.49
12000	74.79	75.29	74.79	74.68
24000	75.10	78.37	75.10	75.66
50000	**77.78**	**78.97**	**77.78**	**77.98**

Table 8.2 BERT Performance over different sample sizes

- **RoBERTa Performance**

  Refer to Table 8.3 for the BERT model performance at different sample size. It is clearly visible from the performance table that the best observation for the

RoBERTa model is noted when the sample size is set to 24000 with highest evaluation metrics.

Sample Size	Accuracy	Precision	Recall	F1 Score
6000	70.75	71.83	70.75	70.97
12000	73.33	74.11	73.33	73.32
24000	**77.60**	**78.81**	**77.60**	**77.80**
50000	77.55	79.29	77.55	77.71

Table 8.3 RoBERTa Performance over different sample sizes

8.4.3 Model Comparison

Models were trained and then tested on varying sizes of dataset and then compared to find out the best-performing model.

- **Initial Performance (6000 samples)**: RoBERTa slightly outperforms BERT in F1 Score, indicating better generalization on small datasets.

Metric	BERT	RoBERTa	Winner
Accuracy	70.33	70.75	RoBERTa
Precision	74.22	71.83	BERT
Recall	70.33	70.75	RoBERTa
F1 Score	70.49	70.97	RoBERTa

Table 8.4 Metric-wise comparison for Sample size = 6000

RoBERTa wins (3 out of 4 metrics)

- **Mid-range (12000 samples)**: BERT has a small edge over RoBERTa in all metrics, showing it learns better with moderate data initially.

Metric	BERT	RoBERTa	Winner
Accuracy	74.79	73.33	BERT
Precision	75.29	74.11	BERT
Recall	74.79	73.33	BERT
F1 Score	74.68	73.32	BERT

Table 8.5 Metric-wise comparison for Sample size = 12000

BERT wins (4 out of 4 metrics)

- **Large-scale Training (24000 samples)**: RoBERTa surpasses BERT in all metrics. This marks a significant shift in favor of RoBERTa's training efficiency and deeper contextual understanding.

Metric	BERT	RoBERTa	Winner
Accuracy	75.10	77.60	RoBERTa
Precision	78.37	78.81	RoBERTa
Recall	75.10	77.60	RoBERTa
F1 Score	75.66	77.80	RoBERTa

Table 8.6 Metric-wise comparison for Sample size = 24000

RoBERTa wins (4 out of 4 metrics)

- **Maximum Training (50000 samples)**: BERT slightly edges out RoBERTa in Accuracy, Recall, and F1 Score, but the gains are marginal (F1 Score: 77.98 for BERT vs. 77.71 for RoBERTa).

Metric	BERT	RoBERTa	Winner
Accuracy	77.78	77.55	BERT
Precision	78.97	79.29	RoBERTa
Recall	77.78	77.55	BERT
F1 Score	77.98	77.71	BERT

Table 8.7 Metric-wise comparison for Sample size = 50000

BERT wins (3 out of 4)

8.4.4 Key Observations

The highlighting observations of the model comparison for the selected models are listed below —

- **RoBERTa exhibits more stable growth** in performance with increasing data, especially between 6000 and 24000 samples.
- **BERT peaks slightly higher** at 50000 samples, but the improvement over RoBERTa is minimal and may not justify the extra training and inference cost.
- **F1 Score**, a critical metric in mental health classification, is highest for:
  - BERT at 50000 (77.98)
  - RoBERTa at 24000 (77.80)

The marginal difference suggests **RoBERTa generalizes better with fewer samples** and reaches near-peak performance faster than BERT.

8.5 DEPLOYMENT

In the final phase of the project, the focus shifted to deploying the mental health classification system for real-world use. Among the two models evaluated—BERT and RoBERTa—RoBERTa was selected as the final model due to its superior performance on key evaluation metrics such as accuracy and F1-score. The selected model was integrated into an interactive web application developed using **Flask**, enabling users to input text about their feelings and receive real-time mental health assessments along with therapist recommendations. The application was successfully deployed via **Ngrok** on **Google Colab**, making it publicly accessible without the need for complex external hosting services or infrastructure, thus ensuring ease of use and portability.

8.5.1 Final Model Selection and Justification

The final model for deployment is selected on the basis of the evaluation results and considering proper reasoning to ensure best performance.

- After evaluating both models comprehensively:
    - **RoBERTa trained on 24000 samples** is selected for deployment.
- Justification:
    - It achieves a near-optimal **F1 Score** (77.80), very close to BERT's best (77.98), but with **half the data**.
    - **Lower computational cost** for training and inference than the 50000-sample models.
    - Better **scalability and generalization** for production environments.
    - This makes RoBERTa not only efficient but also robust for real-time mental health prediction.

8.5.2 Model Deployment

It is clear from the outcome of the model comparison that the best model that outperforms is RoBERTa. Its ability to generalize well on unseen data and faster convergence during fine-tuning made it an optimal choice for real-world deployment. During training, the model was saved and zipped in the local system and then integrated in the application.

- **RoBERTa demonstrated superior performance** across nearly all key evaluation metrics when compared to other models, particularly at larger training sample sizes.
- Its extensive pretraining on a diverse and comprehensive corpus enabled it to capture nuanced linguistic features and better interpret user input related to mental health conditions.
- **Based on its high F1 Score and consistent performance**, the RoBERTa model trained on 24,000 samples was selected for final deployment in the mental health prediction application.

The selected RoBERTa model was deployed through a **Flask-based web application**, allowing users to input their emotional state in natural language. The application predicts the underlying mental health condition and recommends appropriate therapist based on the identified issue. For cloud-based accessibility and ease of sharing, the app is hosted using **ngrok** on **Google Colab**, which generates a secure public URL, enabling users to interact with the model in real time without any local installation. This setup ensures portability, simplicity, and a user-friendly interface for mental health assessment.

8.6 APPLICATION INTEGRATION

8.6.1 Application Overview

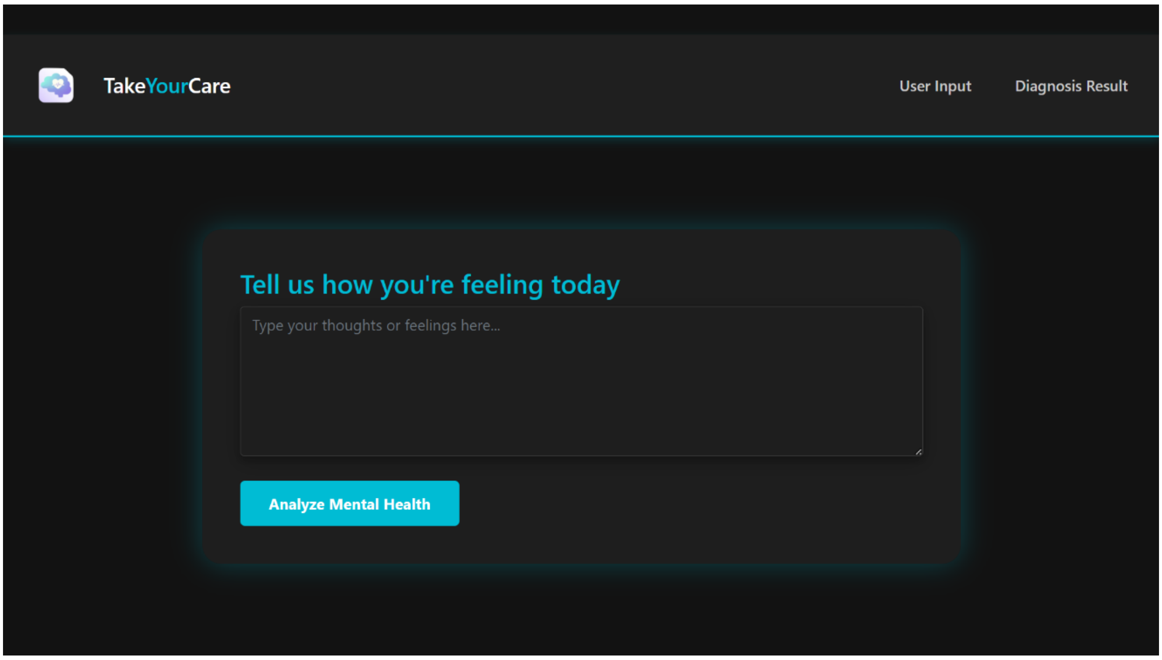

Fig. 8.2 "Take Your Care" App Interface

The application UI in the Figure 8.2 (in Light theme) tells user to enter the text about "*how he feels*?". The inputs and outputs for the application are –

- **Input**: Text input about user's feelings.
- **Output**:
    - Detected Status (can be among "Normal", "Depression", "Anxiety", etc).
    - Therapist Recommendation (depends on the detected status) –
        i. If "**Normal**", no therapist recommendation.
        ii. If other than "**Normal**", therapist recommendation with contact details of therapist.

8.6.2 Technology Used

The technology stack required for the development and deployment of model on the server and implementing in the application is listed below in the Table 8.8 along with tools and technologies with their purposes categorised into the usage purpose such as modeling & training, web deployment, and development tools –

Category	Tool/Technology	Purpose in Project
Modeling & Training	Transformers (Hugging Face)	Fine-tuning RoBERTa for mental health classification.
	Datasets (Hugging Face)	Loading and preprocessing labeled datasets.
	PyTorch	Backend for training and inference.
	Scikit-learn	Accuracy and F1-score computation.
	Pandas & NumPy	Handling data and numerical operations.
	Matplotlib, Seaborn	Visualizations purposes
Web Deployment	Flask	Building and serving the web application.
	Ngrok	Hosting the Flask app online via a public URL on Google Colab.
	Google Colab	GPU-enabled environment for both training and app deployment.
Development Tools	Jupyter Notebook	Interactive coding for model training and experimentation.
	VS Code	Local development of Flask routes and HTML templates.

Table 8.8 Technology used

8.6.1 Application Flow

The flow of application is shown in detail in the Fig. 8.3. We will deeply understand this flow with the steps followed by the user to get the result on his text input while interacting with the application.

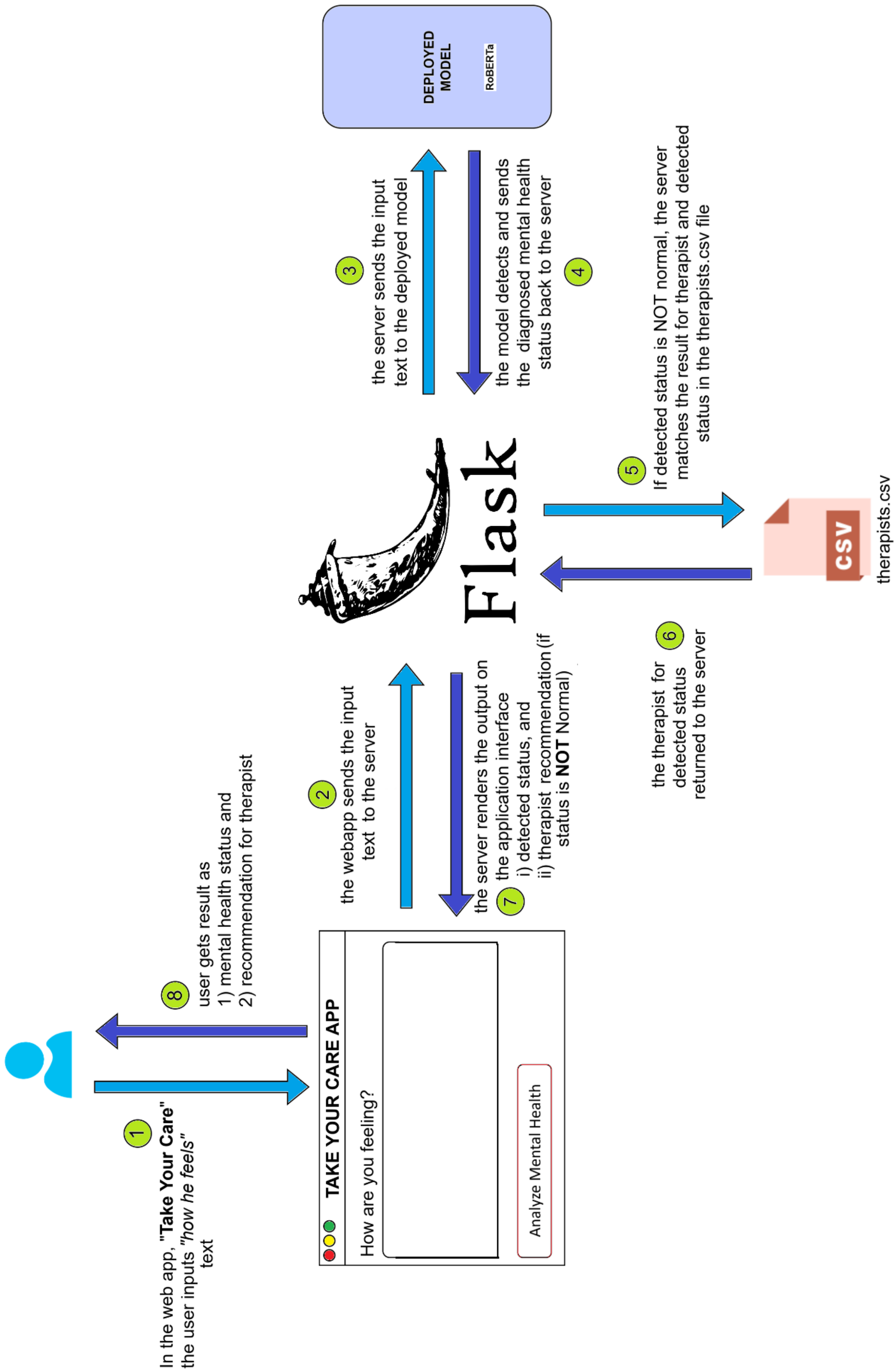

Fig. 8.3 Detailed Flow of Application

The Fig. 8.3 illustrates the user journey for the Mental Health Detection System, highlighting how users engage with the platform and receive AI-generated feedback on their mental health. Let’s break it down step by step:

- **User Visits**
  - The user accesses the application via a public URL exposed by **Ngrok**, which tunnels the **Flask-based web app** running on **Google Colab**.
  - The user sees a text box to describe how they are feeling (it is a **required input**)
- **Application to Flask Server**
  - Upon submitting the form i.e., when the user clicks on the button to detect or analyze the mental health, the **Flask application** receives the user input.
- **Flask Server to Model**
  - The input is forwarded to the deployed mode – **RoBERTa** for mental health status prediction.
- **Model to Flask Server**
  - The model processes the text and returns the **predicted mental health condition** to the Flask server.
- **Flask Server fetches Therapist Recommendation**
  - If the predicted status is **not "Normal"**, the server selects a relevant **therapist recommendation** from the dataset.
- **CSV to Flask Server**
  - The server reads from a CSV file containing therapist data and chooses one appropriate for the detected condition.
- **Server sends back to Application**
  - Flask sends the following back to the frontend –
    - i. Detected Mental Health Status
    - ii. Recommended Therapist (if applicable)
- **Application to user**
  - The application displays –
    - i. Mental Health Detected Status
    - ii. Therapist Recommendation when status is not “**Normal**”

8.6.2 Screenshots

The step-wise application workflow is demonstrated in the following screenshots –

- **Step 1: User Visits the Application**

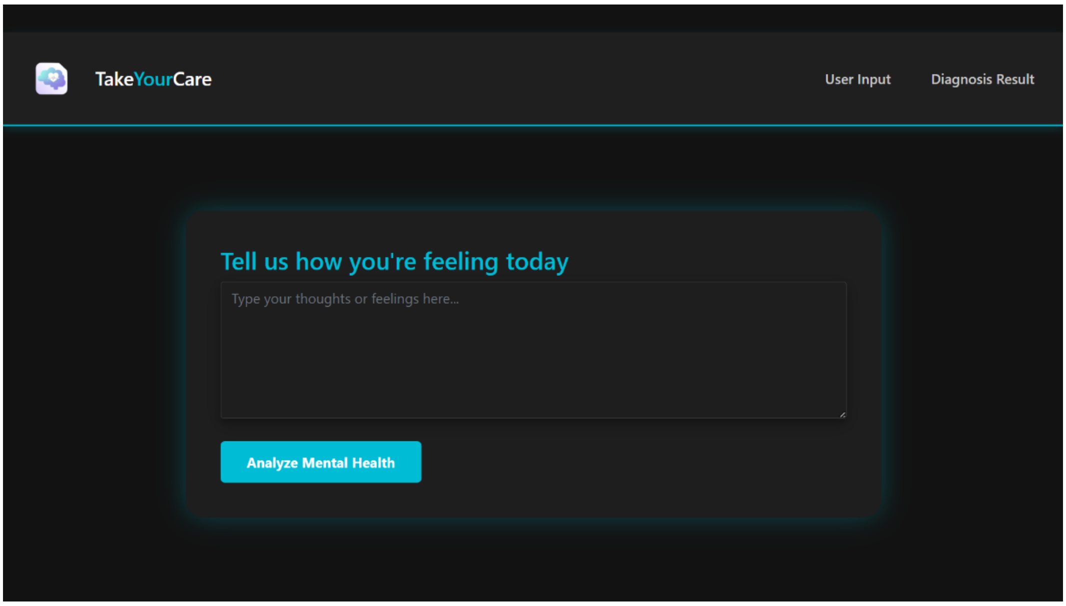

Fig. 8.4 Home page asking for user input

- **Step 2: Inputs the text about "*how he feels?*"**

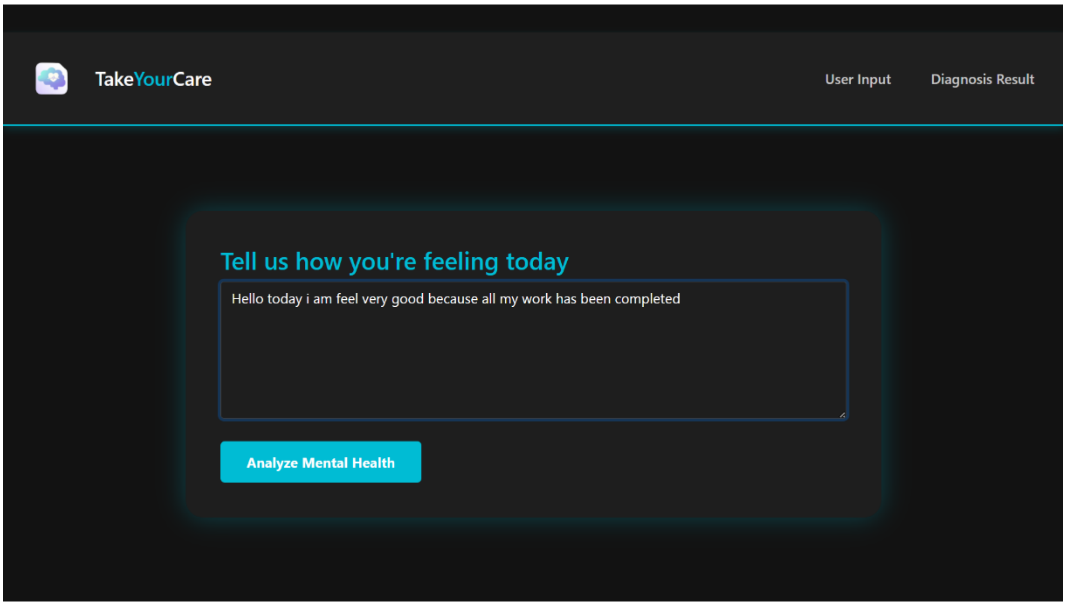

Fig. 8.5 User inputs text about "how he feels?"

➢ **Step 3: Submits the Input, Navigated to the Diagnosis Result Page**

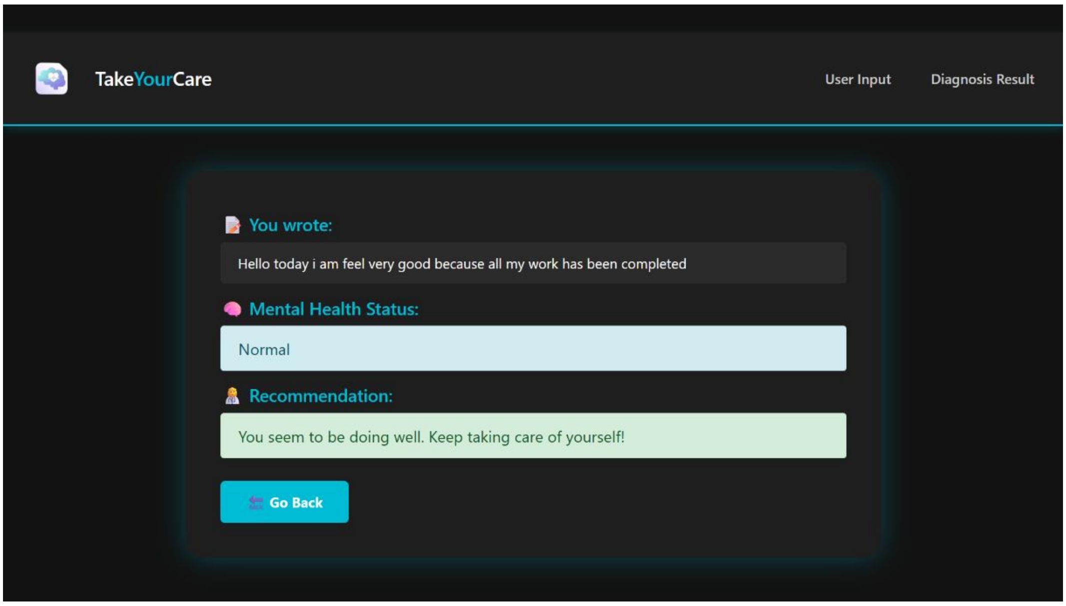

Fig. 8.6 Results page with detected mental status

The screenshots (refer to Fig. 8.4, Fig. 8.5 and Fig. 8.6) above illustrate the basic and complete flow of the mental health prediction application, from user input to the display of diagnosis results and therapist recommendations. These visuals demonstrate how a user interacts with the application by entering their emotional state, how the system processes the input using the deployed RoBERTa model, and how the output is rendered in a user-friendly format. For detailed test cases showcasing predictions across various mental health statuses such as Depression, Anxiety, Suicidal thoughts, and Personality Disorder, please refer to **Chapter 9**.

CHAPTER 9

TESTING

9.1 OVERVIEW

This chapter outlines the testing methodologies applied to evaluate the performance and robustness of both the machine learning model and the deployed application. Testing ensures that the system functions as expected across various scenarios and edge cases.

9.2 MODEL TESTING

The goal of model testing is to evaluate the effectiveness of the trained RoBERTa classifier in correctly predicting mental health conditions from user statements. All the evaluation metrics were utilized to ensure proper testing of the model and the best outperforming model was then selected with following metrics in Table 9.1 below.

Metric	Value
Accuracy	77.60%
Precision	77.81%
Recall	77.60%
F1-Score	77.80%

Table. 9.1 Deployed Model's Evaluation Metrics

9.3 APPLICATION TESTING

9.3.1 Functional Testing

Test Case	Input	Expected Output	Test Case Status
TC1	"I'm feeling low and anxious"	Anxiety	✅ Passed
TC2	"Life is meaningless"	Suicidal	✅ Passed

Table 9.2 Sample Test Cases

Functional testing was carried out using various sample inputs to ensure appropriate condition detection and recommendations. Our application was tested to ensure correct prediction in most of the scenarios. Sample testcases that passed in our application are shown in Table 9.2. Let's see the test cases for all possible detected mental status:

➢ **Test Case 1: Mental Status Detected is "Normal"**

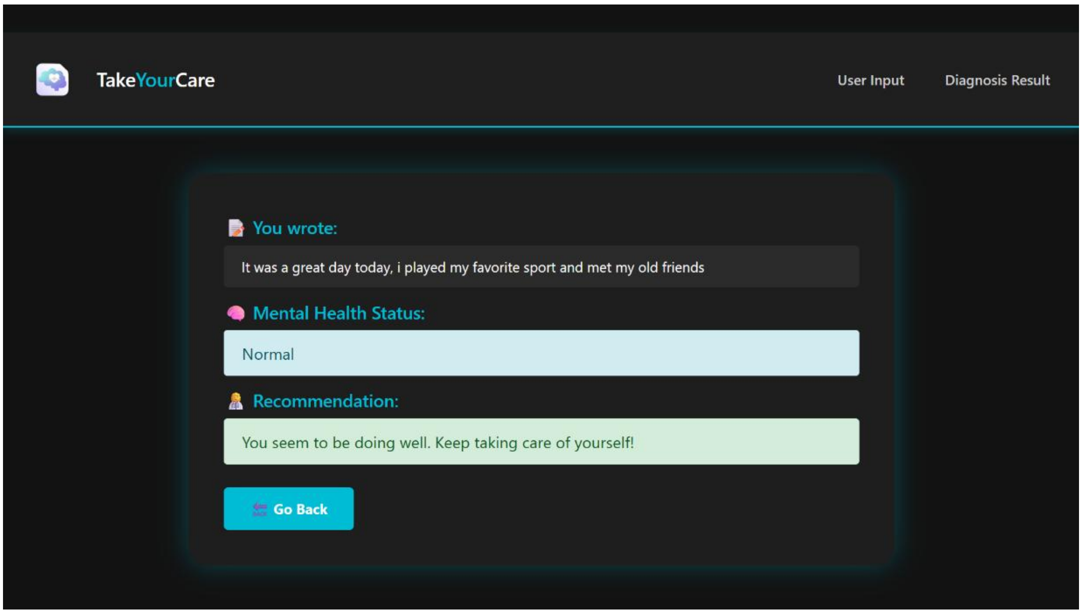

Fig. 9.1 Result Screen for Mental Health Status = "Normal"

- **Input:** *It was a great day today; I played my favorite sport and met my old friends.*
- **Output:** Normal

➢ **Test Case 2: Mental Status Detected is "Anxiety"**

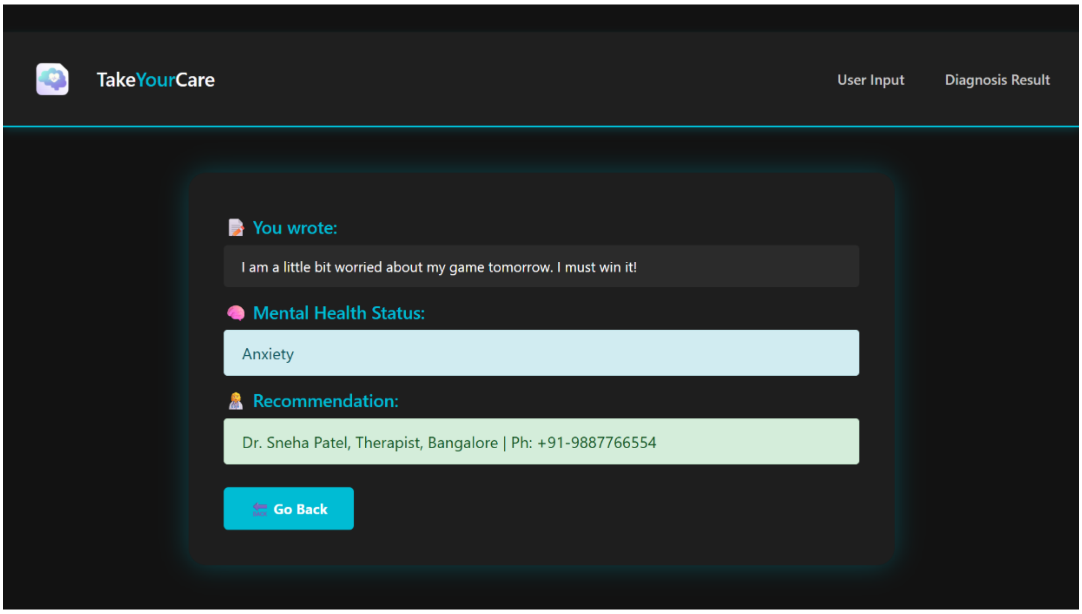

Fig. 9.2 Result Screen for Mental Health Status = "Anxiety"

- **Input:** *I am a little bit worried about my game tomorrow. I must win it!*
- **Output:** Anxiety

➢ **Test Case 3: Mental Status Detected is "Depression"**

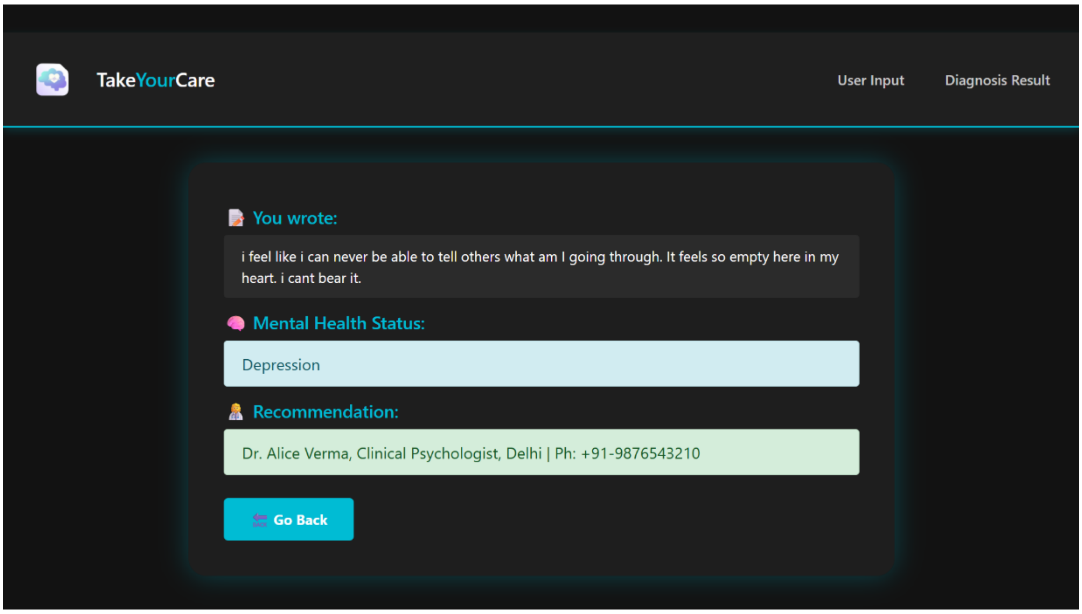

Fig. 9.3 Result Screen for Mental Health Status = "Depression"

- **Input:** *I feel like I can never be able to tell others what am I going through. It feels so empty here in my heart. I cant bear it.*
- **Output:** Depression

➢ **Test Case 4: Mental Status Detected is "Stress"**

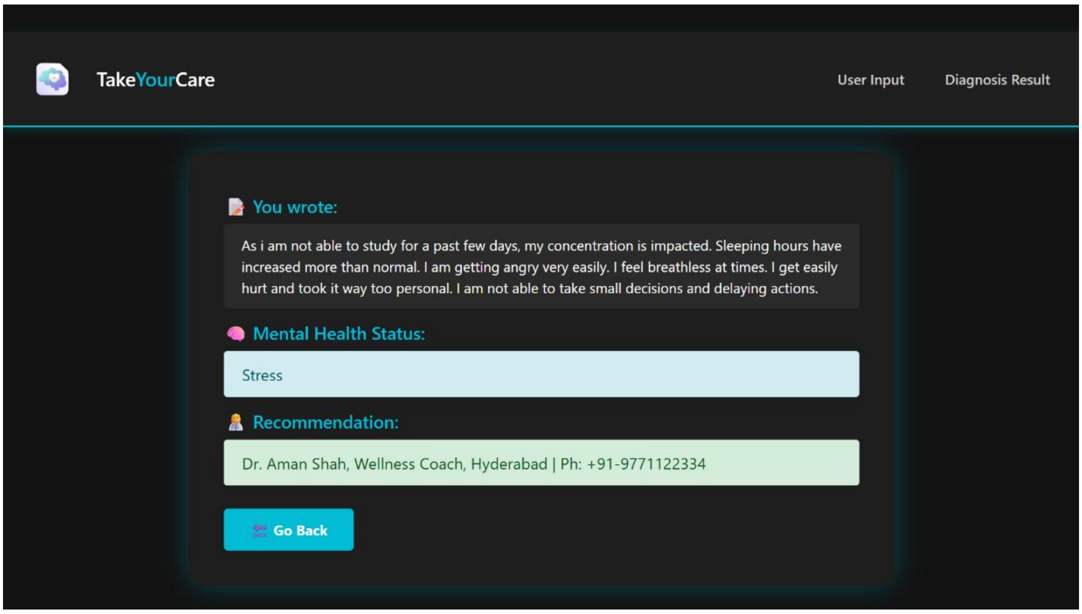

Fig. 9.4 Result Screen for Mental Health Status = "Stress"

- **Input:** *As I am not able to study for a past few days, my concentration is impacted. Sleepings hours have increased more than normal. I am getting angry very easily. I feel breathless at times. I get easily hurt and took it way too personal. I am not able to take small decisions and delaying actions.*
- **Output:** Stress

➢ **Test Case 5: Mental Status Detected is "Bipolar"**

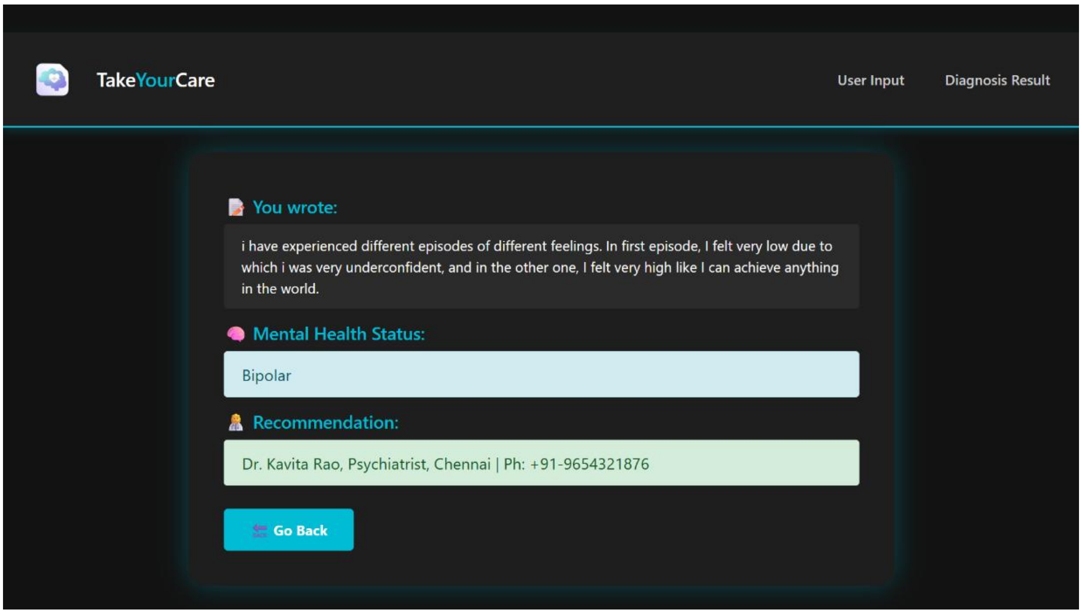

Fig. 9.5 Result Screen for Mental Health Status = "Bipolar"

- **Input:** *I have experienced different episodes of different feelings. In first episode, I felt very low due to which I was very underconfident, and in the other one, I felt very high like I can achieve anything in the world.*
- **Output:** Bipolar

➢ **Test Case 6: Mental Status Detected is "Suicidal"**

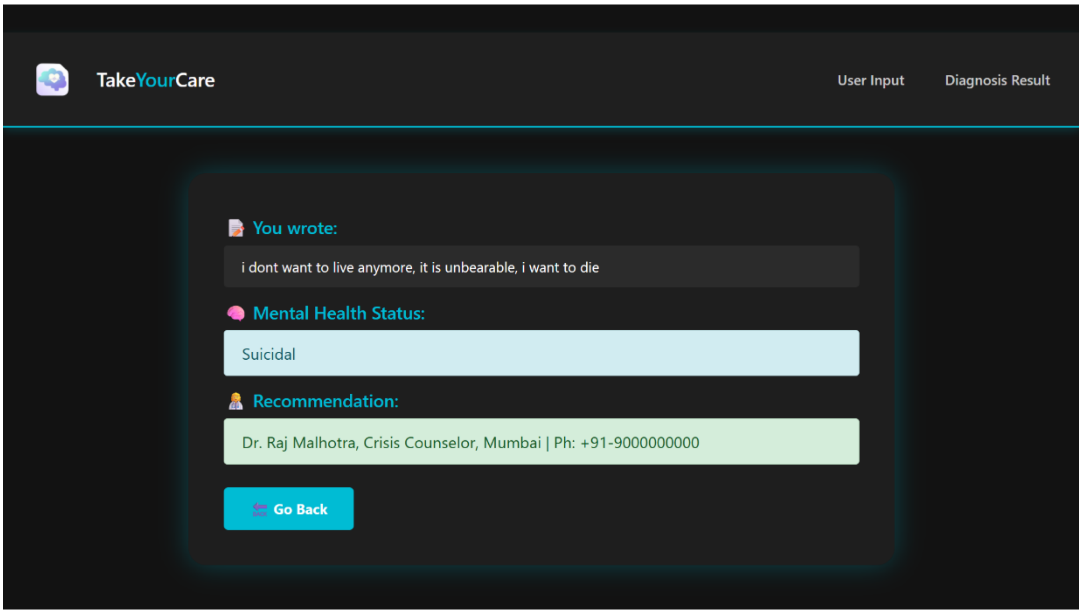

Fig. 9.6 Result Screen for Mental Health Status = "Suicidal"

- **Input:** *I don't want to live anymore, it is unbearable, I want to die*
- **Output:** Suicidal

➢ **Test Case 7: Mental Status Detected is "Personality disorder"**

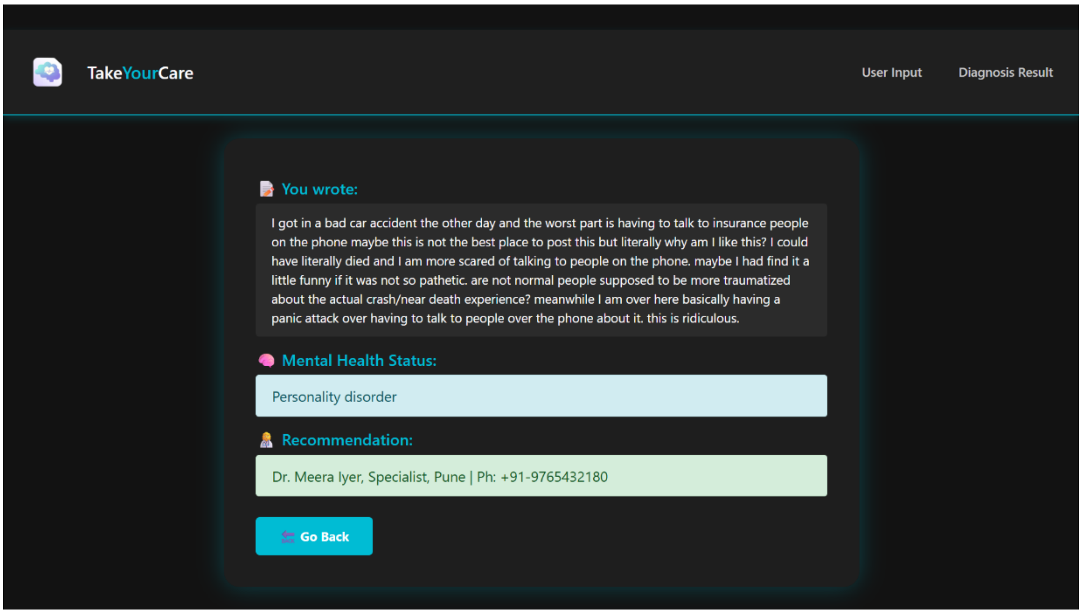

Fig. 9.7 Result Screen for Mental Health Status = "Personality disorder"

- **Input:** *I got in a bad car accident the other day and the worst part is having to talk to insurance people on the phone maybe this is not the best place to post this but literally why am I like this? I could have literally died and I am more scared of talking to people on the phone. Maybe I had find it a little funny if it was not so pathetic. Are not normal people supposed to be more traumatized about the actual crash/near death experience? Meanwhile I am over here basically having a panic attack over having to talk to people over the phone about it. This is ridiculous.*

- **Output:** Personality disorder

9.3.2 UI Testing

Manual testing was performed on the **Flask-based web interface**, and the following UI screenshots represent test cases conducted to verify the functionality of key features:

➢ **User input handling** – Ensuring the system accepts and correctly processes user-submitted text.

- **Case 1**: When the user gives **empty text input** or he/she doesn't enter the text and presses "**Analyze Mental Health**" button, this case is handled and the result screen appears as shown in Fig. 9.8.

- **Case 2**: When the user initially visits the application, he/she might intentionally or unintentionally click on the navbar's *Diagnosis Result* Tab and gets navigated to the results page. The results page in this case appears as shown in Fig. 9.9.

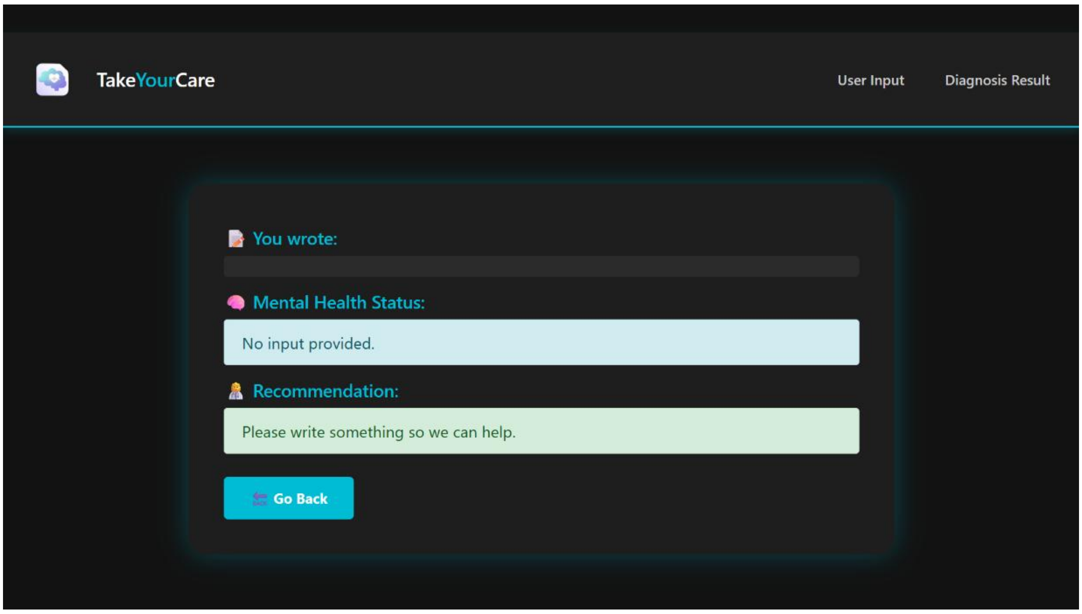

Fig. 9.8 When user tries to detect empty text's mental status

- **No Input**
- **Output:** No input provided.

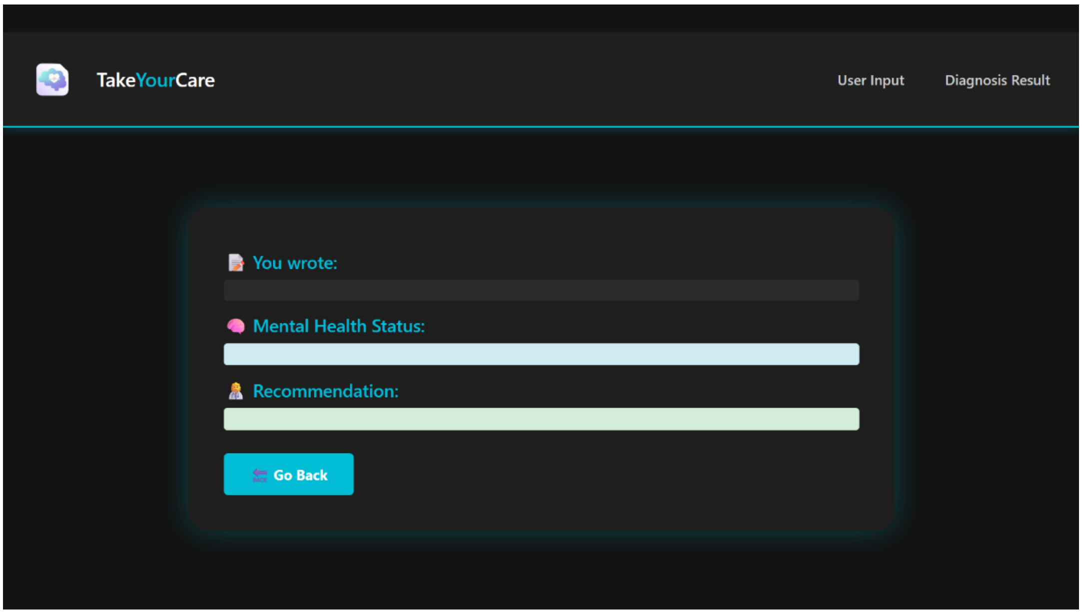

Fig. 9.9 When user visits application and directly jumps to the results page

- **No input**
- **No output**
- No process of detection executed.

➢ **Real-time prediction response** – Validating that the mental health condition is accurately predicted upon submission.

➢ **Therapist recommendation display** – Confirming the system recommends suitable therapists when a condition is detected. In all above test cases discussed, the therapist recommendation is displayed in all cases of mental health statuses except "Normal".

- **Error handling** – Testing scenarios with missing input or invalid requests to ensure graceful error management. The proper error handling is ensured for the application.

9.4 OVERFITTING AND UNDERFITTING ANALYSIS

To evaluate and improve the model's generalization capability:

- **Training vs Validation Loss** was closely monitored across epochs to detect signs of overfitting or underfitting.
- A **class-weighted loss function** was employed to ensure balanced learning across imbalanced classes and to reduce the risk of overfitting on dominant classes.
- **Early stopping** was integrated based on validation performance to prevent overtraining.
- Experiments using varying training subset sizes (e.g., **6,000 to 50,000 samples**) were conducted to simulate different data availability scenarios and assess the model's robustness and stability.

9.5 ERROR ANALYSIS

A sample of misclassified user statements was manually reviewed to understand potential weaknesses in model predictions:

- **Ambiguous inputs** with overlapping symptoms (e.g., anxiety vs. stress) often led to misclassification, suggesting the model struggles with nuanced emotional distinctions.
- **Short or vague statements** were more prone to incorrect predictions due to lack of sufficient context for the model to interpret.

These observations highlight areas for future improvement, such as incorporating **context-aware data augmentation techniques**, **longer input sequences**, or **additional metadata** to improve interpretability and classification accuracy.

9.6 CONCLUSION

The testing phase validated that the system performs effectively in detecting mental health conditions from user-inputted text. Both the RoBERTa-based model and the Flask-based application successfully passed all functional and integration tests. Minor improvement areas—particularly in handling edge cases and ambiguous inputs—were identified. Several of these issues were addressed through enhancements in preprocessing and model input handling to improve overall robustness.

CHAPTER 10

FUTURE SCOPE

The successful implementation of a RoBERTa-based mental health detection and therapist recommendation system marks a significant step toward leveraging artificial intelligence in the mental healthcare domain. However, the scope for enhancement remains vast, both in terms of technical capabilities and real-world application.

10.1 EXPANSION OF DATASET

- **Multilingual Support**: Currently, the model only supports English text. Expanding the dataset to include multiple languages can help reach a more diverse population.
- **Inclusion of Audio/Video Inputs**: Incorporating speech and facial expression analysis can enrich the understanding of user emotions and provide multimodal insights.
- **Larger Clinical Datasets**: Partnering with mental health organizations to include real-world clinical data can help improve the generalizability and clinical accuracy of the model.

10.2 MODEL IMPROVEMENTS

- **Domain-Specific Models**: While RoBERTa has been effective, future work can explore domain-adapted models such as **MentalRoBERTa** or **DistilMentalBERT** for more refined performance.
- **Multi-Label Classification**: Introduce models that can detect comorbid mental health conditions (e.g., stress and anxiety coexisting).
- **Incorporation of Severity Analysis**: Extend the model to not only classify the type of mental health issue but also its severity (mild, moderate, severe), using sentiment and symptom embedding fusion.

10.3 REAL-TIME MONITORING AND FEEDBACK

- **Mobile Integration**: Develop a mobile version of the application to allow continuous check-ins and real-time support.
- **Emotion Tracking Over Time**: Include timeline-based analysis of user responses to track improvement or deterioration.
- **Feedback Loops**: Incorporate user and therapist feedback to continuously improve model predictions.

10.4 ETHICAL AND PRIVACY CONSIDERATIONS

- **Federated Learning**: In the future, federated learning can be adopted to train the model on user data without compromising privacy.

- **Bias Mitigation**: Conduct fairness assessments to reduce potential bias against specific age groups, genders, or regions.

10.5 CONCLUSION

This project lays the groundwork for scalable, AI-driven mental health assessment tools. By addressing limitations and pursuing these future directions, the system has the potential to evolve into a comprehensive digital mental health assistant, capable of assisting both individuals and mental health professionals in early diagnosis, continuous monitoring, and effective treatment planning.

CHAPTER 11

CONCLUSION

- The project successfully developed a deep learning-based system for detecting various mental health conditions using user-generated text data.
- Transformer-based language models, specifically **BERT** and **RoBERTa**, were employed for text classification to identify categories such as Depression, Anxiety, Stress, Bipolar, and others.
- A custom-labeled dataset was preprocessed, tokenized, and structured for fine-tuning transformer models to suit the multi-class classification task.
- Class imbalance was addressed by incorporating **class-weighted loss functions**, improving model fairness across underrepresented mental health classes.
- Four key metrics — **Accuracy, Precision, Recall, and F1 Score** — were used to objectively compare model performance.
- Evaluation results showed that model performance improved significantly with increasing training data, confirming the importance of dataset size in fine-tuning transformer models.
- **RoBERTa outperformed BERT** in terms of F1 Score at a dataset size of 24,000, achieving an optimal balance between precision and recall.
- Models trained on 50,000 samples provided slightly better performance for BERT but at the cost of higher computation and potential overfitting.
- The **RoBERTa model trained on 24,000 samples** offered the best trade-off between performance and efficiency and was chosen for final deployment.
- A scalable and efficient **mental health detection application** was built around the chosen RoBERTa model to enable real-time inference from user inputs.
- The application is capable of recommending professional help (therapist suggestion module) based on the detected condition, adding practical utility.
- Performance insights indicate that domain-specific fine-tuning and well-curated datasets are key to building accurate mental health prediction systems.

- Limitations include reliance on labeled data, possible biases in language use, and lack of contextual metadata (e.g., age, location, background).
- Future work can explore **MentalBERT/MentalRoBERTa**, larger transformer models like **DeBERTa** or **Longformer**, and **multimodal inputs** (e.g., audio + text).
- Integration with therapy chatbots, mood trackers, or social platforms could expand the system into a full-featured mental health support platform.

CHAPTER 12

REFERENCES

1. Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2019). **BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding**. *Proceedings of NAACL-HLT*. arXiv:1810.04805

2. Liu, Y., Ott, M., Goyal, N., Du, J., Joshi, M., Chen, D., ... & Stoyanov, V. (2019). **RoBERTa: A Robustly Optimized BERT Pretraining Approach**. *arXiv preprint arXiv:1907.11692*. arXiv:1907.11692

3. Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., ... & Polosukhin, I. (2017). **Attention is All You Need**. *Advances in Neural Information Processing Systems*, 30.

4. Benton, A., Mitchell, M., & Hovy, D. (2017). **Multitask learning for mental health using social media text**. *EACL*.

5. Guntuku, S. C., Yaden, D. B., Kern, M. L., Ungar, L. H., & Eichstaedt, J. C. (2017). **Detecting depression and mental illness on social media: an integrative review**. *Current Opinion in Behavioral Sciences*, 18, 43–49.

6. Wolf, T., Debut, L., Sanh, V., Chaumond, J., Delangue, C., Moi, A., ... & Rush, A. M. (2020). **Transformers: State-of-the-Art Natural Language Processing**. *EMNLP 2020: Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing: System Demonstrations*.

7. Sarkar, S. (2022). *Sentiment Analysis for Mental Health* [Dataset]. Kaggle. https://www.kaggle.com/datasets/suchintikasarkar/sentiment-analysis-for-mental-health

8. Trotzek, M., Koitka, S., & Friedrich, C. M. (2018). **Utilizing neural networks and linguistic metadata for early detection of depression indications in text sequences**. *IEEE Transactions on Knowledge and Data Engineering*, 32(3), 588–601.

9. Rajput, S., Arora, A., & Kansal, V. (2021). **A survey on sentiment analysis methods, applications, and challenges**. *AI Open*, 2, 68–85.

10. Hugging Face. (2020). **Transformers: State-of-the-art Natural Language Processing for Pytorch and TensorFlow 2.0**. https://huggingface.co/transformers/

11. PyTorch. (2025). **An open-source machine learning framework**. https://pytorch.org/

12. Google Colab. (2025). **Colaboratory: A Research Tool for Machine Learning Education and Research**. https://colab.research.google.com

13. Barros, J. M., Duggan, J., & Rebholz-Schuhmann, D. (2020). **The application of Natural Language Processing to mental illness detection: A narrative review**. *NPJ Digital Medicine*, 3(1), 1–11.

14. Schwartz, H. A., Eichstaedt, J. C., Kern, M. L., Dziurzynski, L., Lucas, R. E., Agrawal, M., ... & Ungar, L. H. (2014). **Towards assessing changes in degree of depression through Facebook**. *Proceedings of the Workshop on Computational Linguistics and Clinical Psychology: From Linguistic Signal to Clinical Reality*.

15. Hossain, M.M., Hossain, M.S., Mridha, M.F. *et al.* **Multi task opinion enhanced hybrid BERT model for mental health analysis**. *Sci Rep* **15**, 3332 (2025). https://doi.org/10.1038/s41598-025-86124-6

16. Simon D’Alfonso, **AI in mental health**, Current Opinion in Psychology, Volume 36, 2020, Pages 112-117, ISSN 2352-250X, https://doi.org/10.1016/j.copsyc.2020.04.005.