A

Project Report

on

Mental Health tracking and surveillance using CBT-based chatbot and BERT classifier in students

submitted as partial fulfillment for the award of

BACHELOR OF TECHNOLOGY DEGREE

SESSION 2024-25

in

Computer Science and Engineering

by

ANANYA VERMA (2100290100028)

ARPITA (2100290100034)

ASHUTOSH VERMA (2100290100036)

Under the supervision of

DR. PREETI GARG

KIET Group of Institutions, Ghaziabad

Affiliated to

Dr. A.P.J. Abdul Kalam Technical University, Lucknow

(Formerly UPTU)

May 2025

DECLARATION

We hereby declare that this submission is our own work and that, to the best of our knowledge and belief, it contains no material previously published or written by another person nor material which to a substantial extent has been accepted for the award of any other degree or diploma of the university or other institute of higher learning, except where due acknowledgment has been made in the text.

Signature

Name: Ashutosh Verma, Ananya Verma, Arpita

Roll No.:2100290100036, 2100290100028, 2100290100034

Date:

CERTIFICATE

This is to certify that Project Report entitled "Mental Health tracking and surveillance using CBT-based chatbot and BERT classifier in students" which is submitted by Ananya Verma, Arpita and Ashutosh Verma in partial fulfillment of the requirement for the award of degree B. Tech. in Department of Computer Science & Engineering of Dr. A.P.J. Abdul Kalam Technical University, Lucknow is a record of the candidates own work carried out by them under my supervision. The matter embodied in this report is original and has not been submitted for the award of any other degree.

.

Dr. Preeti Garg	**Dr. Vineet Sharma**
Associate Professor	**(Dean CSE)**

Date:

ACKNOWLEDGEMENT

It gives us a great sense of pleasure to present the report of the B. Tech Project undertaken during B. Tech. Final Year. We owe special debt of gratitude to Dr. Preeti Garg, Department of Computer Science & Engineering, KIET, Ghaziabad, for her constant support and guidance throughout the course of our work. Her sincerity, thoroughness and perseverance have been a constant source of inspiration for us. It is only her cognizant efforts that our endeavors have seen light of the day.

We also take the opportunity to acknowledge the contribution of Dr. Vineet Sharma, Dean of Computer Science & Engineering, KIET, Ghaziabad, for his full support and assistance during the development of the project. We also do not like to miss the opportunity to acknowledge the contribution of all the faculty members of the department for their kind assistance and cooperation during the development of our project.

We also do not like to miss the opportunity to acknowledge the contribution of all faculty members, especially faculty/industry person/any person, of the department for their kind assistance and cooperation during the development of our project. Last but not the least, we acknowledge our friends for their contribution in the completion of the project.

Date:
Signature:
Name : Ashutosh Verma, Ananya Verma, Arpita
Roll No.: 2100290100036, 2100290100028, 2100290100034

ABSTRACT

Mental health issues are widespread yet often underdiagnosed, particularly among students whose struggles are frequently overlooked. Early detection of mental stress can be pivotal in preventing more severe psychological issues and understanding patterns that lead to chronic conditions. This study focuses on improving student mental health through a novel, user-centric chatbot solution. The chatbot leverages Cognitive Behavioral Therapy (CBT) and Behavioral Activation (BA) techniques to provide empathetic interactions, allowing students to express their concerns and manage stress effectively. For sentiment analysis and stress detection, various machine learning and deep learning models—including Support Vector Machines (SVM), Random Forests, and Recurrent Neural Networks (RNNs)—were employed to classify user input on a scale of 1 (Very Bad) to 5 (Very Good). Among these, BERT demonstrated the highest accuracy of 85.7%, establishing it as the most effective model for this task. Our results indicate that AI-driven approaches offer significant potential for enhancing mental health support among students.

In addition to the strong performance of transformer-based models such as BERT, RoBERTa, DistilBERT, and XLNet, this study highlights the feasibility of deploying mental health tools even in low-resource environments, without the need for extensive hardware or large-scale data. While the current system is focused on student populations, future iterations aim to expand its reach through multilingual support, cultural adaptability, and integration into institutional platforms such as university portals or national e-health services. This work lays a scalable foundation for accessible, intelligent, and empathetic mental health support systems that can extend beyond academia to broader demographics and regions in need.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

BERT	Bidirectional Encoder Representations from Transformers
CBT	Cognitive Behavioral Therapy
NLP	Natural Language Processing
DFFN	Deep Feed-Forward Networks
CNN	Convolutional Neural Networks
RNN	Recurrent Neural Networks
ML	Machine Learning
MERN	MongoDB ExpressJs ReactJs NodeJs
DL	Deep Learning
PTSD	Post-Traumatic Stress Disorder
IEEE	Institute of Electrical and Electronics Engineers
SVM	Support Vector Machine
KNN	K-Nearest Neighbour
LSTM	Long Short-Term Memory
VADER	Valence Aware Dictionary Sentiment Reasoning
AUC	Area Under the Curve
GPs	General Practitioners
GPU	Graphics Processing Unit
CSS	Cascading Style Sheet
API	Application Programming Interface

CHAPTER 1
INTRODUCTION

1.1 INTRODUCTION

1.1.1 Background

There is much more to mental fitness than just the absence of any mental illnesses or problems. Individuals with good mental health can cope with life's challenges, realize their full potential, learn and work efficiently, and contribute to their communities. The ability to make decisions, build relationships, and impact on our environment, both individually and collectively, is vital to health and wellness. Societal and individual development also depends on it

1.1.2 Problem Statement

Tracking of mental health amongst students as their mental state can affect their social as well as academic lives. Good mental health corresponds to a healthy mind, and students who are healthy mentally will be able to learn more effectively and are more likely to reach their greatest potential. In the end, students who are in a positive state of mind are better able to work together, make decisions, and build relationships. These positive effects benefit them individually and as a community as they mature into adults. This makes it important to detect any signs of mental illness at an early age and possibly eliminate them.

1.1.3 Objectives

- To conduct a comparative study on various pre-trained models for mental health analysis.
- To identify key indicators of poor mental health through surveys, user interaction patterns, and daily routines.
- To determine the root causes of deteriorating mental health and suggest actionable solutions.
- To select the most accurate model for classifying mental health status, enabling early detection and support.

1.1.4 Significance of Project

This project aims to provide ML based models most suitable for ascertaining the quality of mental health of students. The parameters for mental analysis may include survey responses, an AI assistant tracker, day-to-day schedule of the individual, etc. It also aims to determine some of the root causes of deteriorated mental health of students and potentially find a solution for the same.

1.2 PROJECT DESCRIPTION

1.2.1 Overview of Saathi - Mental Health tracking and surveillance using CBT-based chatbot and BERT classifier in students

Saathi is a chat-based system aimed at supporting student mental health. It integrates a CBT-based (Cognitive Behavioral Therapy) chatbot and machine learning classifiers to monitor and interpret emotional well-being.

1.2.2 Key Features

- **CBT-based Chatbot:** Offers empathetic support and guides students through structured conversations to manage stress.
- **Sentiment Classification:** Analyzes user input on a 1–5 scale (Very Bad to Very Good) to assess emotional state.
- **Mental Health Insights:** Identifies patterns and potential causes of mental stress through continuous interaction.

1.2.3 Stakeholders

- **Primary Users:** Students seeking mental health support.
- **Secondary Users:** Anyone with mental health concerns.
- **Developers/Researchers:** Teams working on model development and system optimization.

1.2.4 Technology Stack

- **Frontend:** Reactjs
- **Backend:** Flask, NodeJs - Expressjs
- **ML/DL Models:** BERT, ANN
- **NLP Library:** HuggingFace Transformers
- **Database:** MongoDB

1.2.5 Scope and Limitations

- **Scope:** Focused on student communities in academic environments. Involves real-time emotion tracking and feedback through a chatbot interface.
- **Limitations:** Not a substitute for professional therapy; effectiveness depends on user interaction and dataset coverage.

1.2.6 Impact and Benefits

This project demonstrates the potential of AI in early mental health intervention. Comparative analysis revealed that BERT achieved the highest accuracy (85.7%) in sentiment classification, making it the best candidate for text-based emotion recognition. The solution enhances mental health awareness, supports students in distress, and can assist institutions in providing timely care

CHAPTER 2

LITERATURE REVIEW

2.1 Review of Research Paper: A comprehensive review of predictive analytics models for mental illness using machine learning algorithms[2]

2.1.1 Review

This study explores how mental health—encompassing emotional, psychological, and social well-being—affects our thoughts, emotions, and behaviors. It reviews the use of machine learning for early detection of mental illness, focusing on models, algorithms, and data sources like social media, wearable devices, and surveys. The authors propose a comprehensive method for assessing mental health using these data types and introduce a new taxonomy based on five data domains. The paper highlights the current state, benefits, challenges, and future directions of using machine learning in mental health care, aiming to offer faster and more accurate diagnosis and treatment.

2.1.2 Objective

The main objective of the study is to review the current state of machine learning (ML) applications in mental illness detection, focusing on various data modalities. It also aims to propose a comprehensive methodology that integrates social media data, wearable sensor data, and verbal input for more accurate and timely prediction and support in mental health assessment.

2.1.3 Methodology

- Data Collection: The researchers conducted a systematic review using databases like Scopus, IEEE Xplore, Web of Science, and Google Scholar with relevant keywords. They filtered out duplicates, review papers, and non-experimental studies, ultimately selecting 51 papers published between 2018–2024.

- Taxonomy: The paper classifies mental illnesses into depression, anxiety disorders, bipolar disorder, PTSD, schizophrenia, etc., and reviews ML models applied to each.

- ML Models: Supervised, unsupervised, and reinforcement learning approaches were explored.

- Data Modalities: Five key types of data were analyzed — surveys/interviews, social media posts, audio data, sensor/device data, and multimodal data.

- Proposal: A new integrated system combining primary (social media), secondary (wearable/device), and tertiary (verbal/audio) assessment was proposed to offer a holistic ML-powered diagnostic framework.

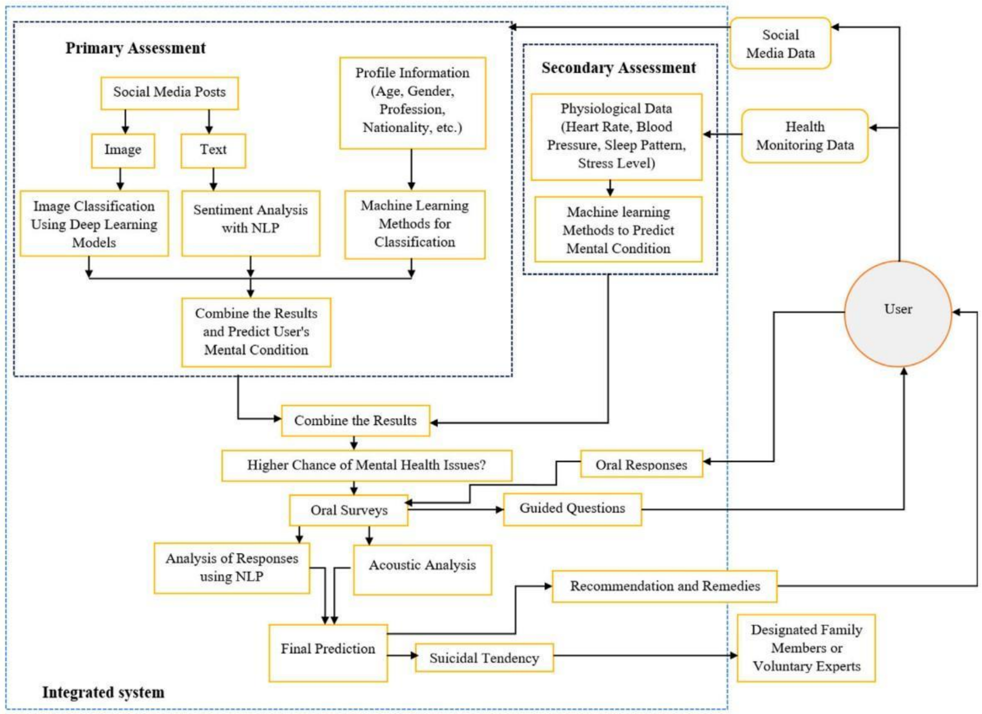

Fig 2.1. Comprehensive Integrated System for Mental Health Assessment.

2.1.4 Key Findings

- Machine learning, especially deep learning and ensemble models, can accurately predict various mental illnesses when applied to the right data types.

- Surveys/interviews and multimodal datasets provided the highest accuracy, while social media data posed challenges due to noise and cultural variance.

- The integration of data from multiple sources enhances prediction performance.

- Personalized and remote mental health monitoring systems are feasible using available technologies.

- A comprehensive system was proposed to facilitate real-time, context-aware mental health assessment incorporating user demographics (age, gender, culture).

2.1.5 Relevance to the Project

This research is directly relevant to any mental health monitoring or predictive system development project. It provides:

- A thorough analysis of what data and ML techniques work best for different mental health issues.
- A validated rationale for using multimodal data (wearables + social media + audio).
- A detailed framework that could be implemented or modified for projects focused on mental health diagnostics or support tools.
- Ethical and accessibility considerations that are crucial for real-world applications.

2.1.6 Limitations

- The review excluded non-English articles, potentially omitting significant findings from countries like China, India, and Saudi Arabia.
- No experimental evaluation of the proposed system was conducted.
- Comparative analysis between different datasets or model performances was not deeply explored.
- Data availability across demographics and the subjective nature of mental health posed challenges in generalization.
- Legal and privacy constraints may limit full implementation of the integrated system.

2.1.7 Conclusion

The study demonstrates that ML holds great promise for early detection and treatment of mental illness when combined with diverse, rich datasets. It proposes a future-ready integrated system that leverages various data sources for more accurate and personalized mental health assessments. Although there are limitations, especially in data access and implementation challenges, the research lays a strong foundation for further innovation and practical deployment of AI-powered mental health tools.

2.2 Review of Research Paper: Sentiment Analysis Using Naive Bayes Algorithm of The Data Crawler: Twitter[8]

2.2.1 Objective

The study aims to analyze public sentiment toward the 2019 Indonesian presidential candidates using Twitter data. It specifically evaluates the effectiveness of the Naïve Bayes classification algorithm in detecting and categorizing sentiments as positive or negative and compares its performance with SVM and K-Nearest Neighbor (KNN) methods.

2.2.2 Methodology

- Data Collection: Tweets related to the Indonesian presidential election were collected via Twitter's data crawler from January to May 2019, totaling 443 data points.
- Text Processing:
    - Tokenization: Splitting tweets into meaningful tokens (words).
    - Text Parsing: Cleaning and preparing the data (removing special characters, standardizing cases).
- Classification: Applied the Naïve Bayes algorithm to classify sentiments, then compared it with SVM and KNN using RapidMiner.
- Evaluation: Accuracy, precision, recall, and F1-score were used to measure model performance.

2.2.3 Key Findings

- The Naïve Bayes algorithm outperformed both SVM and KNN in accuracy:
    - Naïve Bayes: 80.90%
    - KNN: 75.58%
    - SVM: 63.99%
- Naïve Bayes achieved better precision for positive sentiment and performed competitively on negative sentiment.
- The approach proved effective for real-time and small-scale sentiment classification.

2.2.4 Relevance to the Project

- This paper is highly relevant for projects involving:
- Social media analysis for political sentiment or brand reputation.

- Natural Language Processing (NLP) and text classification using classical machine learning methods.
- Baseline model comparison between traditional classifiers.
- Real-world applications where labeled sentiment data from social platforms is utilized to monitor public opinion trends.

2.2.5 Limitations

- Limited Dataset Size: Only 443 tweets were analyzed, which may affect generalizability.
- Binary Sentiment Labels: The classification was limited to only positive and negative, excluding neutral or mixed sentiments.
- Language-Specific Results: The analysis focused on Indonesian tweets, making transferability to other languages uncertain.
- Lack of Deep Learning: Only classical ML algorithms were compared, omitting newer models like BERT or LSTM which may yield better results.

2.2.6 Conclusion

The research concludes that Naïve Bayes is a robust and effective method for sentiment classification in small datasets, especially when dealing with Twitter data. Its simplicity, interoperability, and relatively high accuracy make it a strong candidate for real-time sentiment analysis. Future work could expand to other social media platforms (like Facebook, Instagram) and include broader classification categories or modern ML models.

2.3 Review of Research Paper: Chatbot-Delivered Cognitive Behavioral Therapy in Adolescents With Depression and Anxiety During the COVID-19 Pandemic: Feasibility and Acceptability Study[13]

2.3.1 Objective

The primary goal of the study was to assess the feasibility, acceptability, usability, and preliminary effectiveness of a mobile health (mHealth) app featuring a conversational agent that delivers cognitive behavioral therapy (CBT). The target population was adolescents aged 13–17 with moderate depressive symptoms, who were accessing primary care during the COVID-19 pandemic.

2.3.2 Methodology

- Study Design: 12-week pilot randomized controlled trial (RCT).
- Participants: 18 adolescents aged 13–17, recruited from a network of academically affiliated primary care clinics.
- Intervention: Random assignment to either the CBT-based mHealth app group (n=10) or a waitlist control (n=8).
- Primary Outcome Measure: Depression severity at 4 weeks, measured using the 9-item Patient Health Questionnaire (PHQ-9).
- Additional Measures:
  - Usability, feasibility, and acceptability questionnaires (for teens and parents).
  - Qualitative interviews with 13 primary care providers (PCPs) to understand integration challenges.
- Analysis: Quantitative symptom change, qualitative feedback from adolescents, parents, and PCPs.

2.3.3 Key Findings

- PHQ-9 Reduction:
- App group: Mean score decreased by 3.3 points (moderate to mild category).
- Waitlist group: Mean score dropped by 2 points (no category shift).
- Usability & Acceptance:
- High levels of app usability, acceptability, and feasibility were reported by both teens and their guardians.

- Provider Feedback:
- PCPs were supportive of mHealth tools, especially for early intervention in adolescents with mild to moderate symptoms.
- Safety: No adverse events were reported.

2.3.4 Relevance to the Project

This study is highly relevant for:

- Projects focused on digital mental health interventions for youth.
- Exploring CBT delivery via chatbots in a scalable, user-friendly format.
- Integrating mental health apps into primary care settings for timely intervention.
- Addressing mental health crises during or post-pandemic among adolescents.
- Developing or evaluating mHealth tools targeting depression or anxiety in real-world clinical settings.

2.3.5 Limitations

- Small sample size (n=17 analyzed) limits the generalizability and statistical power.
- Homogeneous demographic: Majority White, female, and privately insured—limiting socio-demographic diversity.
- Short duration of follow-up (4 weeks), which restricts understanding of long-term effects.
- Pilot nature: Not designed to establish definitive effectiveness.
- Exclusion of underserved populations: Rural and socioeconomically disadvantaged communities were underrepresented.

2.3.6 Conclusion

The pilot study successfully demonstrated that a CBT-based conversational agent app is feasible, usable, and acceptable for adolescents with moderate depression in primary care settings. While the study could not confirm clinical effectiveness due to its scale, the promising results suggest that further larger, diverse, and long-term trials are warranted. Such future research should focus on broader populations and integration strategies to support wider adoption in pediatric mental health care.

2.4 Review of Research Paper: Text-Based Emotion Recognition in English and Polish for Therapeutic Chatbot.[10]

2.4.1 Review

This paper explores the application of sentiment and emotion recognition in English and Polish texts for use in therapeutic chatbots. It introduces a bilingual dataset (CORTEX), discusses the importance of recognizing emotional states in human-computer interaction, and evaluates several classification models—including Naïve Bayes, SVM, fastText, and BERT—for their effectiveness in detecting sentiment polarity (3-class) and emotion (9-class) in both languages.

2.4.2 Objective

The primary aim is to **create and evaluate a bilingual (English–Polish) emotion-labelled corpus** and use it to develop and compare models for **text-based sentiment and emotion recognition**, ultimately for integration into **therapeutic chatbots** capable of empathetic dialogue.

2.4.3 Methodology

- **Corpus Creation**: Developed a bilingual corpus (CORTEX) using: Empathetic Dialogues and Daily Dialog datasets (for English).

- **Neural machine translation** (Google Translate API) to generate Polish versions.

- **Labelling**: Aggregated emotion labels into 3 sentiment categories (positive, negative, neutral) and 9 emotion classes.

- **Models Used**:

  - Classical ML: Naïve Bayes (NB), Support Vector Machines (SVM).
  - Shallow NN: fastText.
  - Deep Learning: BERT (fine-tuned separately for English and Polish).

- **Evaluation Metrics**: Accuracy, F1-score, Wilson score interval for statistical confidence.

2.4.4 Key Findings

- **BERT outperformed all models**, achieving:
  - ~93.7% accuracy for English sentiment classification (3-class).
  - ~78.9% accuracy for English emotion classification (9-class).
  - Polish models underperformed English, largely due to MT limitations.
- **High accuracy** in distinguishing neutral vs emotional text (F1 ~97% for neutral).
- Models had **difficulty with ambiguous or overlapping emotions** like “other positives.”

- BERT’s performance gap between English and Polish suggested a need for better native resources and fine-tuning in Polish.

2.4.5 Relevance to the Project

This study is valuable for:

- Designing **emotion-aware therapeutic chatbots**, especially in multilingual contexts.
- Building or using **custom emotional corpora** where low-resource language support is limited.
- Understanding the **impact of translation quality** and language structure on NLP performance.
- Demonstrating that **BERT-based models provide superior performance** and could be baseline models in similar systems.

2.4.6 Limitations

- **Translation quality**: About 10% of Polish translations had minor errors; some affected emotion classification.
- **Label Noise**: Emotion labels from prompts may not fully capture emotional nuance in dialogues.
- **Language gap**: Polish corpora yielded **inferior results**, possibly due to grammar complexity or pretrained embedding quality.
- **Neutral class imbalance**: Many utterances were labelled “neutral,” affecting emotion distribution and model balance.

2.4.7 Conclusion

The study successfully developed a bilingual corpus and validated it through classification experiments. **BERT was the most effective model**, proving viable for emotion-aware dialogue systems. Despite translation-related limitations, the project **set a benchmark** for sentiment and emotion recognition in Polish, encouraging future efforts to expand resources and improve chatbot empathy across languages. The CORTEX corpus is publicly available for further research.

2.5 Review of Research Paper: Sentiment analysis on the impact of coronavirus in social life using the BERT model[12]

2.5.1 Review

This research analyzes public sentiment during the COVID-19 pandemic using Twitter data from both global and India-specific datasets. It employs a hybrid sentiment analysis approach combining rule-based and machine learning techniques, with a strong focus on implementing the **BERT model** for emotion classification. The study aims to assess the emotional impact of the pandemic across different regions using intensity analysis, polarity/subjectivity metrics, and word clouds.

2.5.2 Objective

The study's primary goal is to **perform sentiment and emotion classification** on tweets related to the coronavirus outbreak using **Natural Language Processing (NLP)** and **BERT**. It aims to understand the **social impact of COVID-19**, identify **positive, neutral, and negative sentiments**, and validate how public emotions vary globally and in India.

2.5.3 Methodology

- **Data Collection**:
    - Scraped tweets using Twitter APIs and Tweepy from Jan 20 to April 25, 2020.
    - Created two datasets: one global and one India-specific (filtered using keywords "India" and "Modi").
    - Total of ~597,000 tweets collected.
- **Preprocessing**:
    - Cleaned tweets by removing irrelevant features, links, hashtags, and emojis.
    - Selected key features like tweet text, likes, retweets, and timestamps using **mRMR**.
- **Sentiment Analysis Techniques**:
    - **VADER** for intensity scoring.
    - **TextBlob** for polarity and subjectivity analysis.
    - **WordCloud** for visual representation of common words.
- **Modeling**:
    - Used **BERT** for multi-class emotion classification.
    - Fine-tuned using Hugging Face + PyTorch, with training-validation pipeline.
    - Evaluation via validation accuracy and MCC (Matthews Correlation Coefficient).

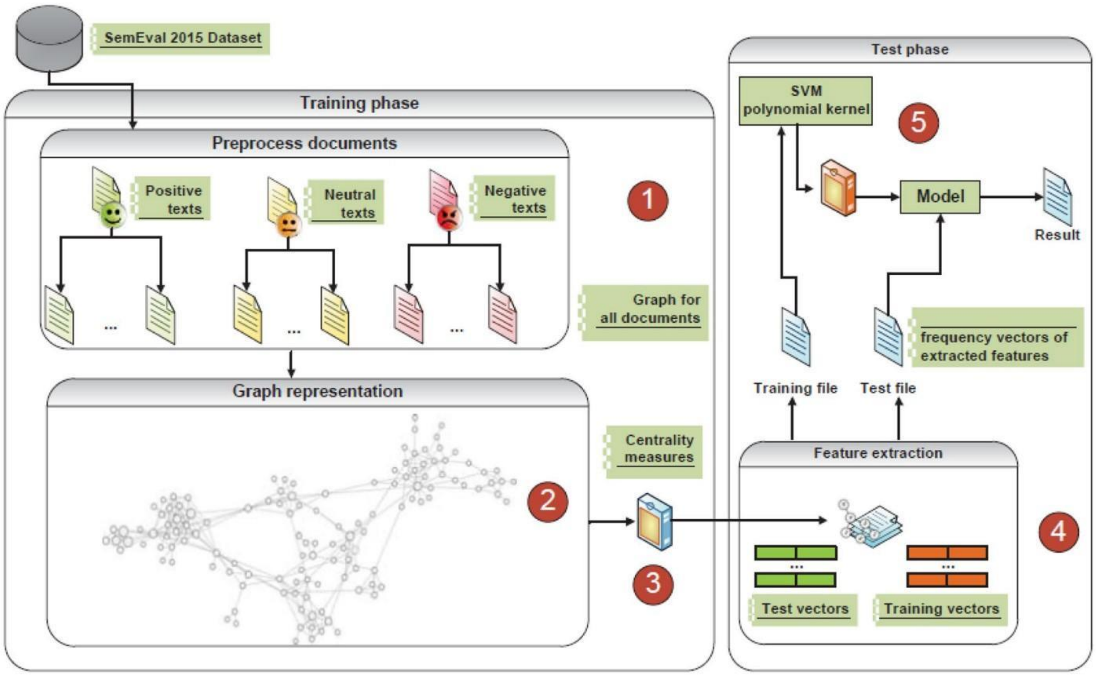

Fig. 2.2 Internal working of sentiment analysis system (Castillo et al. 2015)

2.5.4 Key Findings

- **Accuracy**: BERT achieved a **validation accuracy of ~94%** on emotion classification.
- **Intensity Results**:
  - Majority of tweets were **neutral**, but Indian tweets had **more strong opinions** (both strongly positive and negative).
- **Polarity & Subjectivity**:
  - Indian tweets showed **more emotional variation** with stronger polarity swings.
  - Global tweets remained mostly mildly emotional or neutral.
- **WordClouds** revealed:
  - Common expressions like "lockdown", "virus", "fear", "vaccine", “help”, and "death".
  - Indian users posted **more positive terms** reflecting support for government measures.
- Peaks in engagement (likes/retweets) corresponded with major pandemic events (lockdowns, evacuation plans, etc.).

2.5.5 Relevance to the Project

This paper is relevant to:

- Projects involving **social media mining** for real-time public sentiment tracking.
- Studies using **transformer-based models like BERT** for emotion detection.
- Applications analyzing **geo-specific emotional trends** during global events.
- Integrating **multi-phase sentiment workflows** (scraping, feature selection, modeling) into data pipelines.

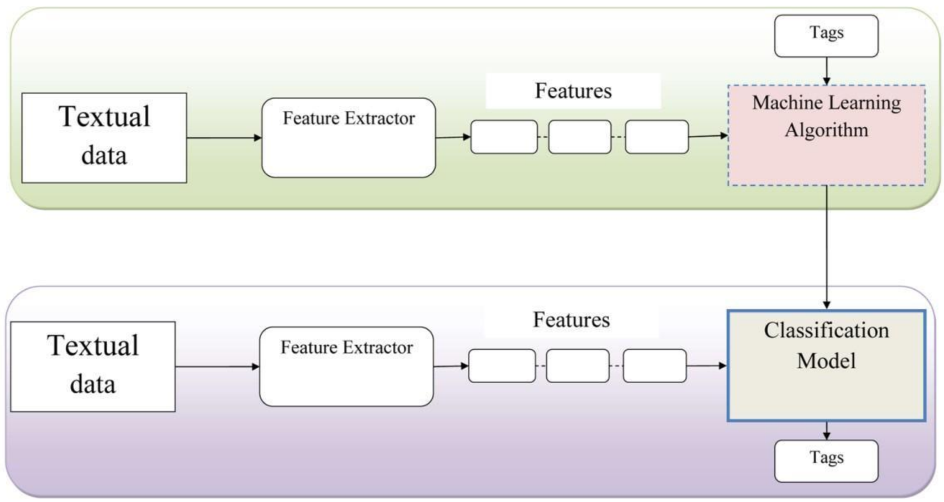

Fig. 2.3 Working of the proposed sentiment analysis model

2.5.6 Limitations

- **Language Constraint**: Only English tweets were analyzed, limiting regional linguistic diversity.
- **Keyword Filtering**: Used basic keyword matching ("India", "Modi") instead of geolocation.
- **Lack of Real-Time Deployment**: Analysis was retrospective, not integrated into a live monitoring system.
- **Imbalanced Data**: Strong skew toward global tweets vs. Indian tweets may affect comparative conclusions.
- **No Human Annotation**: Used pre-labeled data from GitHub without manual validation.

2.5.7 Conclusion

The study effectively demonstrates the utility of BERT and NLP in assessing the emotional and social impact of COVID-19 using Twitter data. It provides detailed insights into how sentiment varied over time and across regions, highlighting India's comparatively more optimistic tone. With a **94% accuracy**, the model proved robust for emotion classification, paving the way for further applications in **crisis communication monitoring** and **policy response evaluation** using social media data.

2.6 Review of Research Paper: A Mental Health Chatbot with Cognitive Skills for Personalised Behavioural Activation and Remote Health Monitoring[11]

2.6.1 Review

This paper presents the design, development, and participatory evaluation of **“Bunji”**, an intelligent mental health chatbot. Unlike traditional CBT-based chatbots, Bunji is grounded in **Behavioural Activation (BA)** therapy, offering **personalised emotional support**, **recurrent engagement**, and **remote health monitoring**. The study aims to demonstrate the feasibility of integrating BA with **AI and NLP techniques** in a mobile chatbot to support individuals experiencing depression and anxiety.

2.6.2 Objective

The main objectives of this study were to:

1. Propose a **conceptual framework** for applying Behavioural Activation (BA) in a chatbot setting.
2. Develop a **cross-platform AI-powered chatbot** with personalisation, emotion tracking, and support features.
3. Evaluate its feasibility and effectiveness through **participatory user trials**, focusing on mood improvement, engagement, and mental health monitoring.

2.6.3 Methodology

- **Conceptual Framework**:
    - Developed through a 3-phase process involving literature review, data mining from mental health forums using the PRIME framework, and expert validation.
- **Chatbot Design**:
    - Implemented using **Rasa**, **BERT**, **DeepMoji**, and **Flair** for NLP and emotion analysis.
    - Features included **personalised conversation**, **activity scheduling**, **gratitude journaling**, **mood tracking**, and **PHQ2/PHQ9 integration**.
    - Remote monitoring via **mood calendars** and **rolling mood averages**.
- **Technical Stack**:
    - Backend: Python, Flask, Redis, Celery, MSSQL.
    - Frontend: React Native for Android/iOS.
    - Hosted on Azure with containerized services.
- **Pilot Study**:

- Duration: 8 weeks.
- Participants: 34 users (out of 318 global downloads) who actively engaged with the app.
- Evaluated across three studies:
    - Mood improvement (pre-post mood score).
    - Impact of personalized engagement.
    - User feedback and qualitative insights.

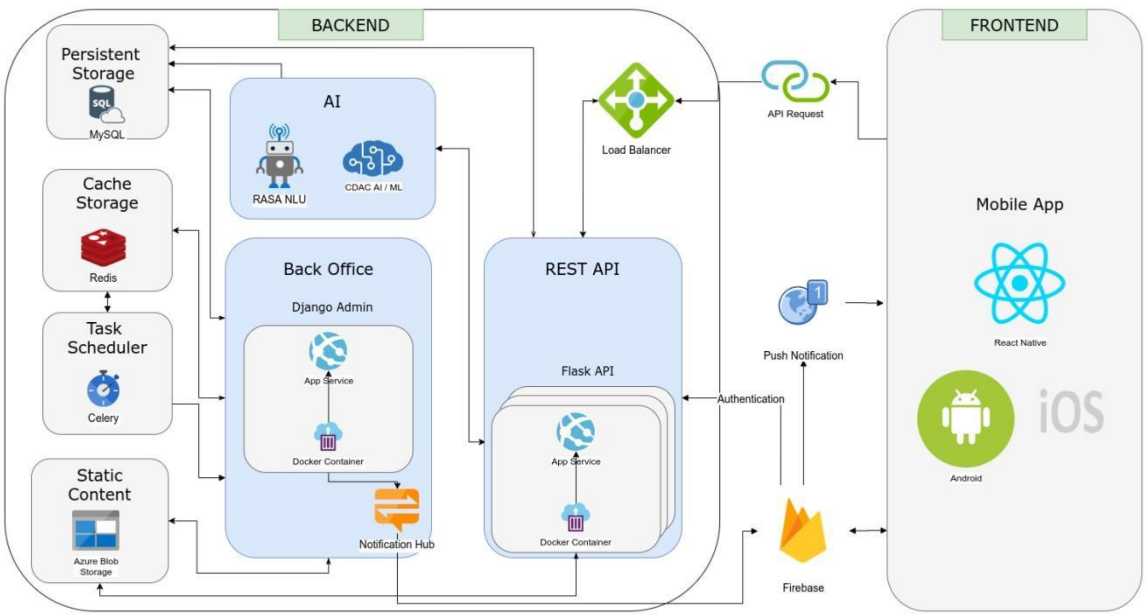

Fig. 2.4 Technical architecture of the BA-based AI chatbot, which is implemented as a smartphone application on both Android and iOS operating systems.

2.6.4 Key Findings

- **Mood Improvement**:
    - Mean mood scores improved from **5.79 to 7.38** (p=0.03), indicating significant positive emotional change.
- **User Engagement**:
    - 88% scheduled activities; 85% used inspiration features; 44% used the gratitude journal.
    - Emotion tracking showed frequent transitions towards **joy, love, and optimism**.
- **Remote Monitoring**:

- Users could monitor mood patterns via calendar and rolling averages.
- Alerts for critical phrases (e.g., self-harm) triggered ethical safeguards.

- **Positive Feedback**:
  - Users appreciated mood tracking, emotional journaling, and activity planning features.
  - Reported increased self-awareness and willingness to share mood data with clinicians.

2.6.5 Relevance to the Project

This work is highly relevant for:

- Projects focused on **AI-powered therapeutic tools** using **Behavioral Activation**.
- Mental health chatbot development with **emotion recognition** and **personalization**.
- Remote health applications require **real-time monitoring** and **mood tracking**.
- Demonstrating the impact of **empathetic conversational agents** in reducing depressive symptoms.

Fig. 2.5 Temporal mood score and mood calendar enabling remote mental health monitoring.

2.6.6 Limitations

- **Short pilot duration** (8 weeks) with limited follow-up.
- **Sample bias**: Most participants were from Australia; global generalization is limited.
- **No clinical validation** against therapists or longitudinal studies.
- Requires further testing on **diverse socio-demographic populations**.
- Integration with human support systems (e.g., therapists, GPs) not evaluated.

2.6.7 Conclusion

The study successfully introduces a novel **BA-based AI chatbot** that combines emotional intelligence, personalized interactions, and remote monitoring to support mental health. Results show promising mood improvements and high usability, suggesting this solution could scale in mental healthcare settings. Future work will explore **community building, gamification**, **expanded activity banks**, and **long-term evaluation** for broader adoption.

2.7 Review of Research Paper : Mental Health Prediction Using Machine Learning: Taxonomy, Applications, and Challenges - by Jetli Chung and Jason Teo[7]

2.7.1 Review

This comprehensive **systematic literature review** analyzes the current landscape of machine learning (ML) techniques used for **predicting mental health disorders** such as schizophrenia, bipolar disorder, PTSD, anxiety, depression, and children's mental health. The authors propose a **taxonomy** for classifying ML methods, data types, and disorder types and evaluate the **performance and applicability** of ML in clinical mental health settings. The review is grounded in the **PRISMA methodology**, reviewing 30 selected papers across multiple ML approaches.

2.7.2 Objective

The key objectives of the paper are to:

- Systematically review ML techniques used to **predict and diagnose mental health problems**.
- Categorize these methods by **type of mental illness and ML model**.
- Analyze their performance (accuracy, sensitivity, specificity, AUC).
- Discuss **existing challenges, gaps**, and propose **future research directions** in applying ML to mental health prediction

2.7.3 Methodology

- **PRISMA Protocol**: Followed for systematic review.
- **Sources**: Peer-reviewed journals, IEEE, Springer, ScienceDirect.
- **Selection**: 142 articles screened, 30 included after filtering.
- **Classification Criteria**:
  - Types of mental disorders: Schizophrenia, Depression & Anxiety, Bipolar Disorder, PTSD, Children's Mental Health.

- ML techniques: Supervised, Unsupervised, Ensemble, Neural Networks, Deep Learning.
- Data types: Questionnaires, clinical interviews, neuroimaging, audio/text, physiological signals.

- **Performance Metrics**: Accuracy, AUC, F1-score, sensitivity, specificity.

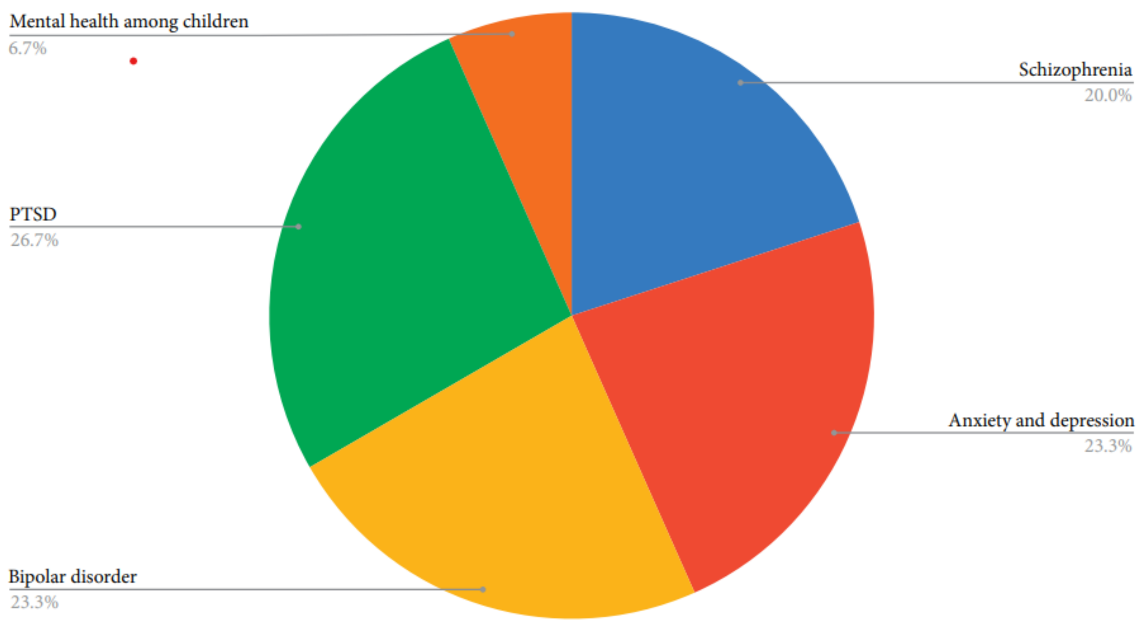

Fig 2.6 Percentage for the number of papers reviewed for each mental health problem

2.7.4 Key Findings

- **Random Forest and Support Vector Machines (SVM)** were most frequently used and showed strong performance.
- **Deep Learning models** like CNNs and Deep Belief Networks showed highest accuracies (up to **96%**) in some cases.
- **Mental health disorder-specific observations**:
    - **Schizophrenia**: Deep learning achieved 94.4% accuracy; other models had varied results (~60-82%).

- **Depression & Anxiety**: CNNs, SVM, and RF showed excellent prediction accuracy (up to **96.8%**).
- **Bipolar Disorder**: Moderate to high accuracy (64–96%); eye-tracking and ECG features used innovatively.
- **PTSD**: RF and Gradient Boosting performed well (AUC up to **0.954**).
- **Children's Mental Health**: Random forest, SVM, and neural nets achieved AUC ~0.74.

- **Data diversity**: Text, voice, neuroimaging, ECG, and questionnaire-based data enriched classification models.

2.7.5 Relevance to the Project

This paper is valuable for:

- Projects involving **predictive modeling of mental disorders using ML**.
- Building data pipelines using multimodal data for **diagnosis support systems**.
- Understanding the comparative strengths of ML models for **specific disorders**.
- Offering a **benchmark dataset and model-performance table** to guide algorithm selection.
- Designing interventions using **clinical and non-clinical data sources**.

2.7.6 Limitations

- **Dataset size limitations**: Many studies used small sample sizes (<100 subjects), reducing generalizability.
- **Diversity of data sources**: Lack of consistency in data types and preprocessing methods across studies.
- **Lack of real-world deployment**: Most models evaluated in experimental setups, not integrated into live systems.

- **Limited cross-cultural representation**: Many studies are localized and may not account for cultural/ethnic variation in mental health symptoms.

- **Comparative difficulty**: Performance comparison across different datasets and disorders is challenging due to inconsistent evaluation standards.

2.7.7 Conclusion

This systematic review confirms that **ML techniques are promising tools** for mental health prediction across diverse disorders. While **Random Forest, SVM, and Deep Learning** models deliver high accuracy in various contexts, real-world application remains limited. Future research should focus on:

- **Standardized, large-scale datasets**,

- **Model generalizability across demographics**, and

- **Ethical considerations** for real-world mental health diagnostics.
  The paper provides a robust **taxonomic foundation** and calls for collaborative, interdisciplinary research to enhance AI-driven mental healthcare.

LITERATURE REVIEW SUMMARY

Table 2.1. A detailed analysis of similar Literary works.

Ref	Objective	Technique/ Algorithm	Findings	Summary
[7]	Overview, critical analysis, and systematic literature on machine learning methods for identifying, diagnosing, and predicting mental health issues.	Multi model deep learning architecture and multiple pre-trained language models.	Top 3 Mean F1-scores for text features – Random forest: 0.73 Gaussian process classification: 0.71 Support vector machine: 0.72 Top 3 algorithms with highest accuracy – Deep Learning: 94.44% Random Forest: 83.33% Logistic Regression: 82.77%	This study reviews 33 ML and DL articles for diagnosing mental health disorders, noting that deep learning shows potential for diagnosing multiple disorders and analysing visual data.
[8]	Public sentiment analysis on Twitter regarding the 2019 Indonesian presidential candidates.	K-Nearest Neighbor (K-NN), Naïve Bayes, Support Vector Machines	Jokowi-Ma'ruf Amin pair: +45.45%, -54.55% Prabowo-Sandiaga pair: +44.32%, -55.68% Mean Accuracy: 80.90%	The K-NN, Naive Bayes and SVM approaches were compared in this study. RapidMiner was used to test the methods, producing highest accuracy in Naïve Bayes Model.
[9]	Using Machine Learning techniques for Sentiment Analysis of Malayalam Tweets.	● Naive Bayes Classifier ● SVM ● Random Forest	The Random Forest classifier achieved the maximum accuracy of 95.6% when used with Unigram and Sentiwordnet, including negation words.	This work classifies Malayalam tweets with four feature sets using NB, SVM, and RF, finding that Unigram with SentiWordNet and negations improved accuracy for all classifiers.
[10]	Emotion recognition for Polish and English texts	● Support Vector Machines ● Naïve Bayes ● BERT ● fastText	The BERT Model performed better than the other algorithms. F1-score: 97.6% (English) and 96.8% (Polish) Sentiment Accuracy: 90% Emotion Accuracy: 80%	The F1-scores and accuracy for Polish were found to be somewhat lower than those for English, with BERT showing the largest disparity.

[11]	Personalized Cognitive Behavioral Therapy through chatbot-as-a-conversational-agent	• Neural Network Model • k-Nearest-Neighbor (kNN)	Before Usage – Total count of users (N):34 Mean mood score: 5.79 Median mood score: 6.50 Shapiro–Wilk statistic (p-value): 0.80 ($p = 3.15 \times 10^{-5}$) After Usage – Total count of users (N):34 Mean mood score: 7.38 Median mood score: 8.00 Shapiro–Wilk statistic (p-value): 0.78 ($p = 1.50 \times 10^{-5}$)	The proposed chatbot uses a neural network and kNN for personalized emotional support and mental health monitoring, with an emergency detector for critical situations, proving effective in various scenarios.
[12]	Sentiment Analysis of tweets during Covid outbreak.	BERT	Validation Accuracy: 94% (approx.)	This study used the BERT model to analyse sentiment in global and Indian tweets during COVID-19. It found Indian tweets more positive, indicating the impact of government measures.

CHAPTER 3

PROPOSED METHODOLOGY

3.1 Introduction

The rise in mental health concerns among students has created a demand for effective, accessible, and non-intrusive methods of addressing emotional well-being. Cognitive Behavioral Therapy (CBT) and Behavioral Activation (BA) are two therapeutic techniques that have shown great promise in helping individuals manage negative thoughts and emotions. CBT focuses on identifying and altering negative thought patterns, while BA emphasizes the role of behavior in influencing emotional states, encouraging individuals to engage in positive activities to improve their mood.

The use of chatbots in delivering mental health support is a novel and effective solution. By utilizing Natural Language Processing (NLP) techniques, chatbots can assess user sentiment in real-time, provide tailored responses, and track emotional changes over time. The sentiment analysis process, facilitated by NLP classifiers like transformers (e.g., BERT), helps the chatbot interpret emotional cues from user inputs and adapt its responses to offer appropriate support.

This project explores the development of a mental health tracking application, leveraging CBT and BA through a chatbot interface. The chatbot uses sentiment analysis powered by advanced transformer models to detect and assess users' emotional states, guiding them through therapeutic techniques and providing recommendations based on their current mental health status.

3.2 Project Development Approach

3.2.1 Development Process

3.2.1.1. Conceptualization and Design

The development of the mental health tracking system began with a clear focus on defining the chatbot's functionality and design. The core objective was to apply Cognitive Behavioral Therapy (CBT) and Behavioral Activation (BA) techniques to offer emotional support. This involved outlining how the chatbot would interact with users, provide therapeutic guidance, and help manage emotions by suggesting positive behavioral changes.

- **Cognitive Behavioral Therapy (CBT)**: The chatbot would identify and challenge negative thought patterns.
- **Behavioral Activation (BA)**: It would encourage users to engage in positive activities despite emotional distress.

3.2.1.2. Data Collection and Preprocessing

To train the chatbot effectively, data related to various emotional states was collected. This dataset included textual expressions representing an individual's mood, categorized into five distinct levels (ranging from "Very Bad" to "Very Good"). The dataset came from various sources such as mental health forums, social media posts, and self-reported experiences.

Categorization of Emotional States: The dataset was labeled according to five emotional categories:

Table 3.1 Category Distribution for BERT Classifier

Category	**Description**	**Assigned Score**
Very Bad	Extreme distress and hopelessness	1
Bad	Sadness, frustration, low energy	2
Neutral	Balanced emotions, indifference	3
Good	Contentment, motivation, general cheerfulness	4
Very Good	Euphoria, high satisfaction, intense excitement	5

Preprocessing steps included filtering out irrelevant text, removing duplicate entries, and ensuring consistent category labeling.

3.2.1.3. Model Selection and Training

For sentiment analysis, we used **transformer-based models**, such as BERT-base-uncased and XLNet, which are capable of understanding and interpreting textual inputs in the context of emotional sentiment. These models were fine-tuned to classify text into one of the five emotional categories based on user input.

- **Tokenization and Input Formatting**: The text data was tokenized using BERT's tokenizer to convert user inputs into token sequences, making them compatible with the model.
- **Classification Algorithms:**
    - BERT-base-uncased
    - XLNet-base-cased
    - DistilBERT-base-uncased
    - RoBERTa-base
- **Training Setup**: The data was split into training, validation, and test sets, with the following parameters:
    - Epochs: 20
    - Batch Size: 8
    - Optimizer: AdamW with warm up steps

- Evaluation Strategy: Steps-based evaluation (every 5 steps)

3.2.1.4. Model Evaluation

The performance of the models was evaluated using several metrics, including accuracy, precision, recall, and F1-score. These metrics were crucial in determining how well the models classified emotional states and how accurately the chatbot could interpret user emotions in real-time.

- **Cross-validation**: Models were tested on multiple datasets to ensure robustness and prevent overfitting.
- **Comparison of Models**: Different transformer models (e.g., BERT, XLNet) were compared to selecting the best-performing model for sentiment analysis.

3.2.1.5 Team Structure and Roles

- **Mentor**: Oversees and coordinates between teams.
- **Frontend Developers**: Focus on creating responsive and interactive user interfaces with Vite and Tailwind CSS.
- **Backend Developers**: Handle server-side logic and API endpoints using Node.js and Express.js.
- **Research Paper**: A team to discuss and share learnings from various research already made and prepare a holistic paper on our idea.

3.2.1.6 System Architecture

- **Machine Learning -** BERT provides a pretrained language classifier so most of the task for NLP is sorted out by it, so the classification model could be trained on a standard machine with GPU acceleration.

  VS code setup with Jupyter Notebook provides an interactive python development environment. Google Collab also provided a great help to test many functionalities stand alone.

- **Full Stack Development -** A full stack architecture, MERN is used to create a platform for serving the ML models along with UI/UX.

3.3.1 High-Level Architecture Overview

The system consists of a user-facing chatbot interface, integrated with a robust backend. The core functionality includes sentiment analysis using NLP models like BERT, which are trained to detect emotional cues from user input. Based on the user's emotional state, the chatbot provides personalized responses using Cognitive Behavioral Therapy (CBT) and Behavioral Activation (BA) techniques to guide the user towards positive actions and mindset.

The architecture is divided into several layers:

- **Frontend Layer**: The ReactJS-based interface provides an interactive space for users to communicate with the chatbot. It displays emotional feedback, progress tracking, and recommended actions.
- **Backend Layer**: A Node.js server handles user data and authentication providing a connection to database, and also specific services to frontend. A Flask app is used to serve two specific APIs for ML models.
- **Data Layer**: The application stores user conversations and emotional states in a secure, encrypted database.

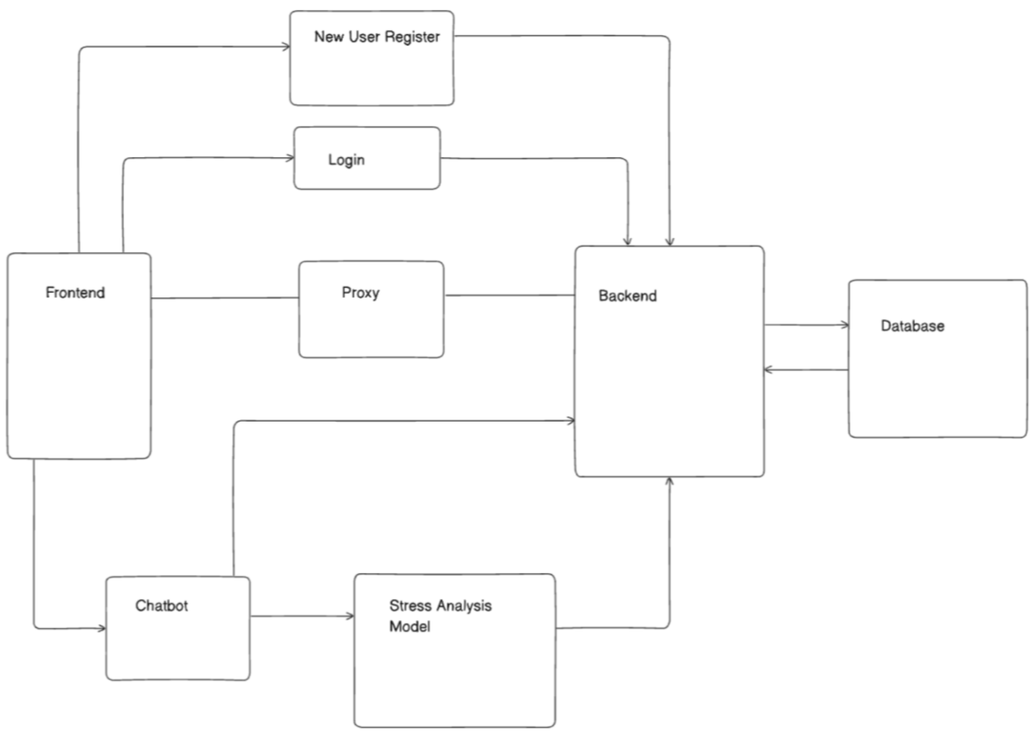

Fig 3.1 Flow Chart for Model Implementation

3.3.2 Data Flow

The data flow involves multiple stages, from user interaction to emotional sentiment classification and the chatbot's response generation. When a user inputs a message, the following steps occur:

1. **User Input**: The user's emotional expression is captured through text.
2. **Sentiment Classification**: The backend utilizes NLP models (such as BERT) to analyze the input, classify the sentiment, and identify the user's emotional state.
3. **Response Generation**: Based on the classified emotional state, the chatbot generates a personalized response, incorporating CBT and BA techniques.

4. **Recommendation**: The chatbot may suggest actions to improve the user's mood, such as engaging in relaxation exercises, breaking tasks into manageable steps, or reflecting on positive experiences.
5. **Feedback Loop**: The system tracks user progress, adjusting recommendations based on ongoing interactions.

3.4 Key Features

The key features of the mental health tracking chatbot include:

- **Emotion Detection**: Real-time analysis of user sentiment using NLP models like BERT and XLNet to detect various emotional states.
- **CBT and BA Integration**: Personalized recommendations based on Cognitive Behavioral Therapy (CBT) and Behavioral Activation (BA) principles to help users manage negative emotions and improve their mood.
- **Sympathetic Conversations**: The chatbot engages in supportive and empathetic dialogues with users, offering emotional consolation and encouraging positive behavior.
- **Privacy and Security**: All communications are end-to-end encrypted with Multi-Factor Authentication (MFA), ensuring user data privacy.

3.5 Technology Stack

3.5.1 Frontend

- **ReactJS**: Used for building the user interface, providing a responsive and dynamic environment for user interactions with the chatbot.
- **Recharts -** A Nodejs library equipped with data visualization tools, used to show analytics made on user data after classification done by BERT-classifier.
- **Tailwind CSS**: A utility-first CSS framework integrated for designing responsive web pages and components. Tailwind ensures a consistent design language across the platform while optimizing performance across screen sizes.

3.5.2 Backend

Nodejs - Used as the backend framework, facilitating server-client interactions, and handling API requests to the sentiment analysis models. It provides various packages and tools for that, from which ExpressJs is used to create the server.

Flask - Additional API endpoint server to server the chatbot and BERT classifier as ML models to the main application.

Python-based Models - Sentiment analysis and emotion detection are powered by Python-based machine learning model, including pre-trained language models of BERT

3.5.3 Database

The application uses MongoDB atlas as a cloud based flexible database. MongoDB provides a document-based NoSQL architecture database, which provides us flexibility in creating and managing the data models.

MongoDB provides security against attacks like SQL injections, and a user-friendly platform to view and make queries in the database.

3.5.4 Hosting and Deployment

- **Frontend on Vercel**: Vercel offers seamless deployment, automatic scaling, and optimized performance for React-based applications.
- **Backend on AWS EC2:** Deployed Server on EC2 instance, which serves the frontend with Elastic IP, which provides a fixed IP address that could be assigned to any EC2 instance. Right now a single instance is capable of serving all the loads.

3.6 Implementation Plan (Only Features)

The features will be implemented in the following phases:

- **Phase 1**: Frontend and backend architecture setup, basic chatbot interface, and integration with sentiment analysis models.
- **Phase 2**: Implement personalized CBT and BA-based recommendations.
- **Phase 3**: Testing and refinement of the emotion detection model and chatbot responses.
- **Phase 4**: Deployment and ongoing monitoring.

3.7 Testing Methodology

- **Unit Testing**: Each module (frontend, backend, sentiment analysis) will be tested independently.
- **Integration Testing**: The entire system will be tested for smooth communication between the frontend, backend, and models.
- **User Acceptance Testing (UAT)**: Real users will test the system in different scenarios to ensure the chatbot provides meaningful and empathetic responses.

3.8 Security Measures

3.8.1 Data Protection

The system will ensure strict adherence to data protection laws, including anonymization and pseudonymization of user data, and end-to-end encryption (E2EE) of all communications.

Users' personally identifiable information (PII) will be fully protected against unauthorized access.

3.8.2 Fraud Prevention

Multiple layers of security, including Multi-Factor Authentication (MFA), will be implemented to prevent unauthorized access to the system. The platform will monitor unusual activity and respond quickly to potential security threats.

CHAPTER 4

RESULTS AND DISCUSSION

4.1 Results

To assess the performance of different transformer-based models for sentiment classification, we evaluated them using a consistent dataset and standard training configuration. Table 3 summarizes the key evaluation metrics—accuracy, F1 score, evaluation loss, precision, and recall—for each algorithm tested.

Table 4.1. Comparison of result parameters of different models

Classifiers	Evaluation Metrics (%)				
	Accuracy	**F1 Score**	**Evaluation Loss**	**Precision**	**Evaluation Recall**
BERT-base-uncased	85.7	86.4	61.1	86.8	86.6
Roberta-base	78.5	77.1	11.1	83.3	77.4
Distilbert-base-uncased	85.7	86.4	59.2	87.1	86.3
XLNet-base-cased	69.0	70.0	138.1	70.8	69.5

While accuracy provides an overall view of the model's correctness, F1 score is a more balanced and meaningful measure in our case, as it combines both precision and recall. BERT-base-uncased and DistilBERT outperformed the other models in both accuracy and F1 score, making them the preferred options for chatbot integration.

All models were fine-tuned on a standard laptop without GPU acceleration. This made the training process significantly slower, but still feasible given the small dataset size of 200 labeled samples. However, for larger-scale implementations, hardware acceleration or cloud-based training environments (e.g., AWS SageMaker, Google AI Platform) would be essential for scaling. Additionally, DistilBERT, being a lightweight version of BERT, showed impressive performance while offering better computational efficiency—ideal for resource-constrained deployments.

4.2 Discussion

4.2.1 Achievements and Impact

The successful fine-tuning of BERT and DistilBERT on a modest hardware setup validated the feasibility of deploying transformer-based sentiment analysis even in low-resource environments. During initial evaluations, the chatbot showed strong potential in providing responsive and context-aware mental health support.

Furthermore, the model's application was tested in a controlled pilot involving a diverse group of consenting students. The trial included monitoring application usage and collecting logs while maintaining full data anonymity and privacy. Post-trial surveys and user interviews assessed the chatbot's ability to demonstrate empathy, supportiveness, and critical thinking. Early feedback showed promising results, with users reporting positive mental health outcomes attributed to chatbot interactions.

4.2.2 Challenges Faced

A key limitation in the development phase was the **limited dataset size (200 samples)**. While it enabled model training on consumer-grade machines, it restricted model generalizability and required careful handling to prevent overfitting. Additionally, training on a **non-GPU setup** significantly increased computation time, creating delays in iteration and experimentation.

Other challenges included:

- Interpreting ambiguous user inputs
- Balancing emotional tone in automated responses
- Ensuring user engagement while avoiding excessive intrusiveness

4.2.3 Comparison with Existing Systems

Unlike many mental health chatbot applications that use rule-based or keyword-driven systems, our solution leverages transformer-based sentiment analysis to derive context and nuance in emotional expression. This allows for more dynamic, personalized, and empathetic responses, aligned with CBT and BA principles.
Additionally, while commercial tools may require expensive subscriptions and internet bandwidth, our lightweight model is optimized for resource-constrained environments, which is especially relevant in educational institutions and rural setups.

4.2.4 Potential Improvements

To enhance scalability and performance, the following improvements are recommended:

- **Dataset expansion**: Collecting larger and more diverse data samples will improve generalization and accuracy.

- **Cloud deployment**: Hosting model inference on cloud platforms like AWS or Azure will accelerate training and real-time response capabilities.
- **Multilingual support**: Incorporating language diversity will help cater to a broader student demographic across India.
- **Speech input support**: Future versions could incorporate voice-to-text capabilities to enable accessibility for visually impaired users.

4.2.5 Societal and Environmental Relevance

The solution addresses a critical societal challenge—access to mental health support among students. In countries like India, where the student-to-counselor ratio is significantly low, such AI-driven solutions can bridge the gap affordably and scalably. By integrating the chatbot into college portals and student apps, it can become a first-line support system, helping students manage mental health before issues escalate.

Environmentally, by operating in cloud or low-energy computing environments and avoiding high-power GPU usage where possible (especially with DistilBERT), the application promotes energy-efficient AI development practices.

4.3 Visual Representation

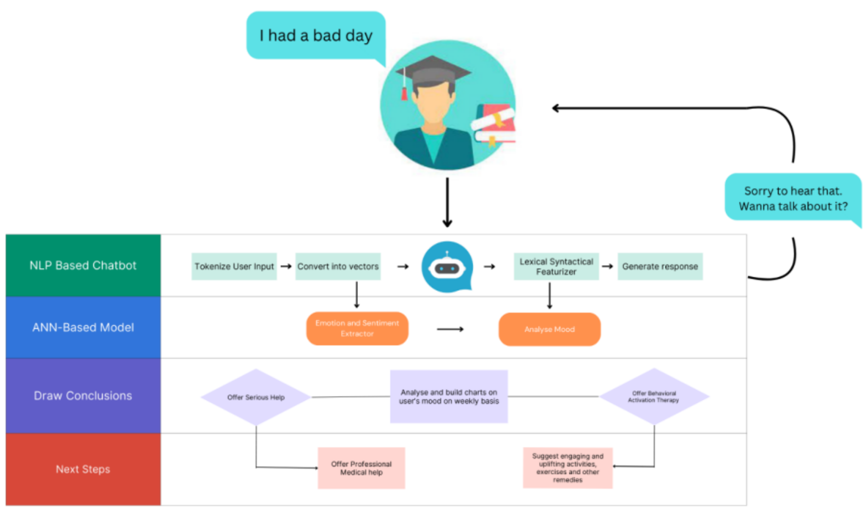

Fig. 4.1 Workings of Saathi- A mental health tracking application.

4.4 Project Screenshots and User Guidance

Home page - A simple and user-friendly interface to direct to chatbot.

Authentication - User need to authenticate first, so a user account is created in database to store user data and further records in a secure way

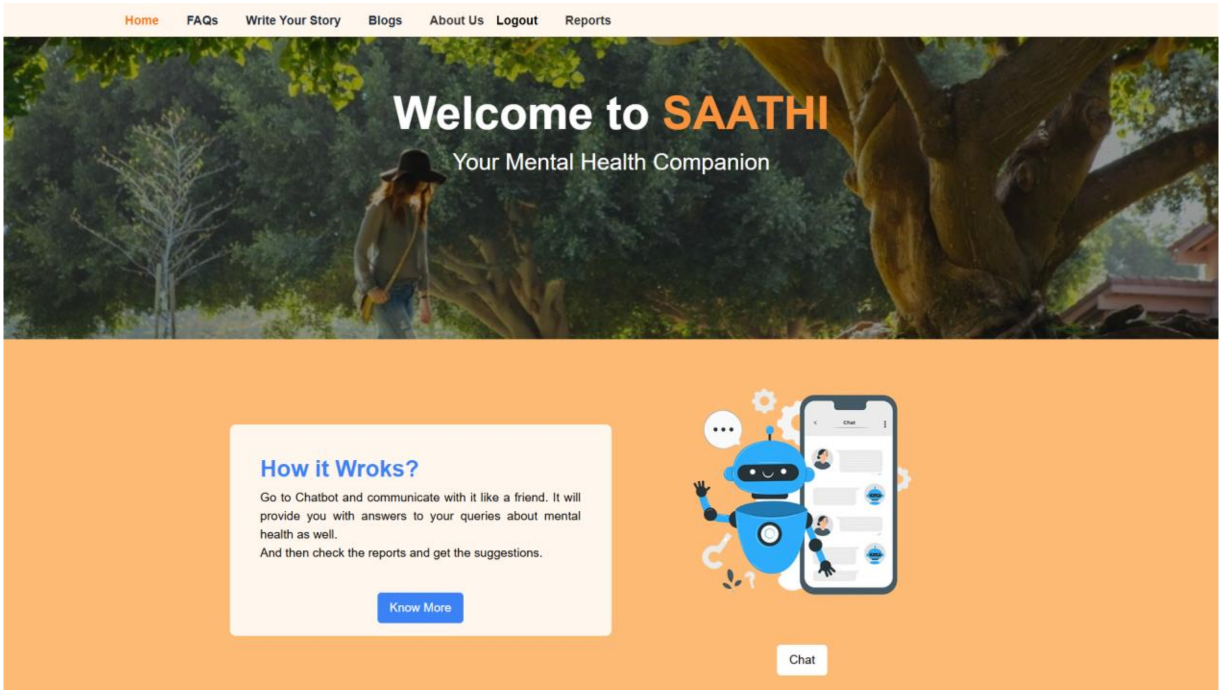

Fig 4.2 Home Page of Saathi - our web application

Chat Interface - A simple user - model chat interface without any distractions to serve users with a compatible environment to communicate to the chatbot and open up with what comes in his mind.

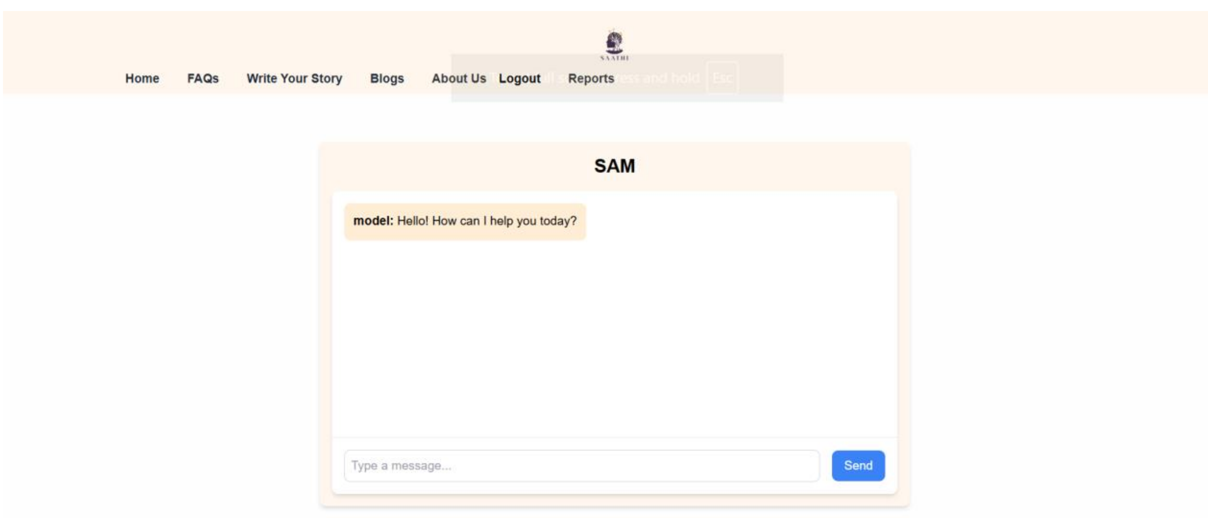

Fig. 4.3 Chat Interface with a default message from Model

Reports and analysis - A detailed Mental Health Report is shown to users on the Reports Dashboards with various metrics and methods, it shows in a holistic way all previous records and also focuses on the last 24 hours as well.

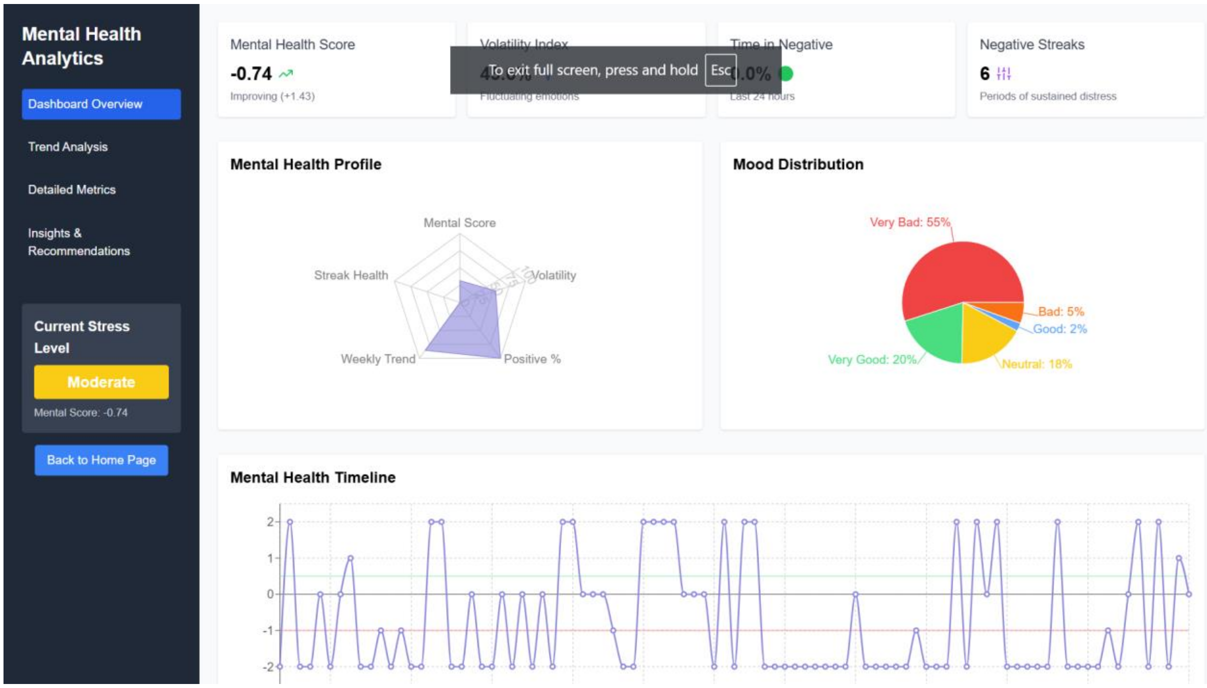

Fig. 4.4 User Dashboard for Mental stress reports analytics and representation

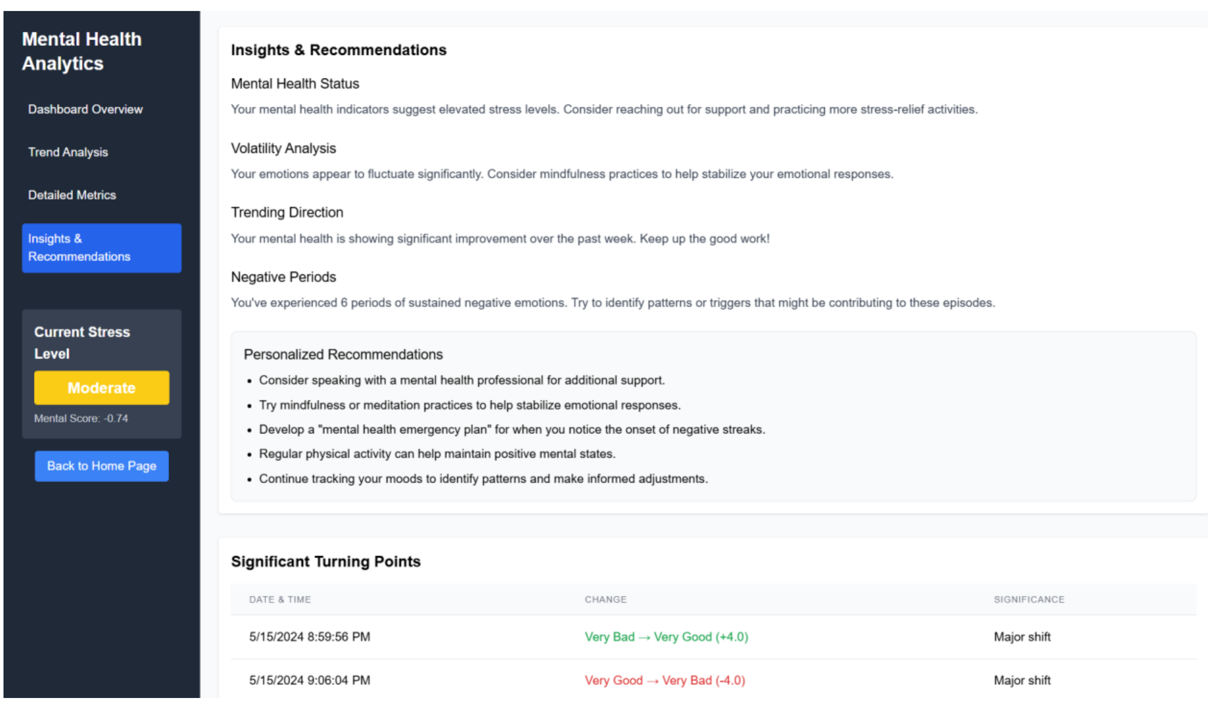

Fig. 4.5 Insights and Recommendation for user auto generated depending upon stress score

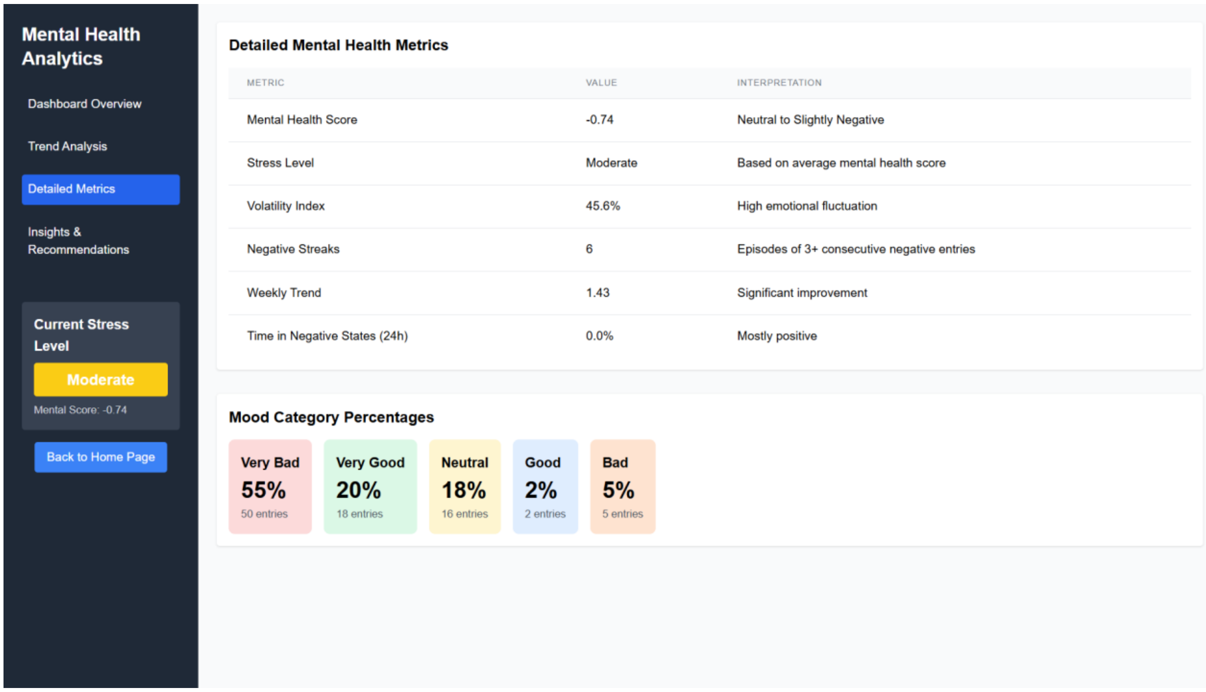

Fig 4.6 Detailed Mental Health metrics showing how user responded in chats

CHAPTER 5

CONCLUSION AND FUTURE SCOPE

5.1 Conclusion

Mental health among students remains an under-addressed area in both educational and medical fields. This project aimed to automate the detection and support of mental well-being through the integration of Cognitive Behavioral Therapy (CBT) with transformer-based sentiment analysis models. The chatbot system—backed by advanced NLP models such as BERT, RoBERTa, DistilBERT, and XLNet—was developed to interact with students, assess their emotional state, and provide therapeutic guidance.

Among the models tested, BERT-base-uncased showed the highest performance with an accuracy of 85.7% and an F1 score of 86.4%, indicating strong reliability in detecting emotional sentiment. The system, even when trained on a modest dataset and without GPU acceleration, demonstrated the feasibility of deploying intelligent mental health tools in low-resource environments.

While current model capabilities are limited by the size and diversity of the dataset, the foundation has been successfully laid for an AI-powered support system that is empathetic, accessible, and scalable.

5.2 Future Scope

Several avenues exist to expand and improve the current system:

- **Dataset Expansion**: Incorporating more diverse data through surveys, chatbot conversations, and regional inputs will improve model accuracy and generalization.
- **Multilingual and Cultural Adaptation**: As mental health expressions vary across cultures and languages, regional customization of training data and models will be necessary for wider adoption.
- **Scalability and Integration**: The system can be scaled and embedded into university portals, mental health platforms, or government-supported e-health initiatives for broader access.
- **Cloud-Based Infrastructure**: Transitioning to scalable cloud services like AWS SageMaker or Google Cloud AI will significantly enhance training, inference speed, and deployment capabilities.
- **Broader Demographic Application**: The chatbot’s use can be expanded to cover working professionals, adolescents, and elderly users, with context-specific behavioral modules.

5.3 Final Thoughts

This project demonstrates a step forward in the fusion of artificial intelligence and mental health support, particularly in student populations. While it is not a replacement for human therapy, the chatbot serves as a first-line, accessible mental health companion—available anytime and capable of triaging emotional distress.

As awareness about mental well-being grows, such AI-powered tools can become essential components of future healthcare ecosystems, particularly in underserved or overstressed environments. With continuous refinement and responsible deployment, this system holds the promise to make mental health support more inclusive, responsive, and proactive.

REFERENCES

1. Kalita, Bhanita: Mental Health of Youth: A Study on Its Importance in Present Scenario. Journal of Clinical Health Research, vol. 13, no. 6, pp. 2299–2303 (2023)

2. Md. Monirul Islam, Shahriar Hassan, Sharmin Akter, Ferdaus Anam Jibon, Md. Sahidullah, A comprehensive review of predictive analytics models for mental illness using machine learning algorithms, Healthcare Analytics, Volume 6, 2024, 100350, ISSN 2772-4425, https://doi.org/10.1016/j.health.2024.100350.

3. Madububambachu U, Ukpebor A, Ihezue U. Machine Learning Techniques to Predict Mental Health Diagnoses: A Systematic Literature Review. Clin Pract Epidemiol Ment Health. 2024 Jul 26;20:e17450179315688. doi: 0.2174/0117450179315688240607052117. PMID: 39355197; PMCID: PMC11443461.

4. Iyortsuun NK, Kim SH, Jhon M, Yang HJ, Pant S. A Review of Machine Learning and Deep Learning Approaches on Mental Health Diagnosis. Healthcare (Basel). 2023 Jan 17;11(3):285. doi: 10.3390/healthcare11030285. PMID: 36766860; PMCID: PMC9914523.

5. Su C, Xu Z, Pathak J, Wang F. Deep learning in mental health outcome research: a scoping review. Transl Psychiatry. 2020 Apr 22;10(1):116. doi: 10.1038/s41398-020-0780-3. PMID: 32532967; PMCID: PMC7293215.

6. Bećirović, Faruk & Smajlhodzic-Deljo, Merima & Salkić, Irma & Softić, Adna. (2024). Assessing the Accuracy of Logistic Regression and Bert in Sentiment Analysis and Mental Disorder Classification. 317-326. 10.5644/PI2024.215.19.

7. Chung, Jetli & Teo, Jason. (2022). Mental Health Prediction Using Machine Learning: Taxonomy, Applications, and Challenges. Applied Computational Intelligence and Soft Computing. 2022. 1-19. 10.1155/2022/9970363.

8. M. Wongkar and A. Angdresey, Sentiment Analysis Using Naive Bayes Algorithm of The Data Crawler: Twitter. 2019 Fourth International Conference on Informatics and Computing (ICIC), Semarang, Indonesia, 2019, pp. 1-5, doi: 10.1109/ICIC47613.2019.8985884.

9. S., Soumya & K.V., Pramod. Sentiment analysis of Malayalam tweets using machine learning techniques. ICT Express. 6. 10.1016/j.icte.2020.04.003 (2020).

10. Zygadło, Artur & Kozłowski, Marek & Janicki, Artur. Text-Based Emotion Recognition in English and Polish for Therapeutic Chatbot. Applied Sciences. 11. 10146. 10.3390/app112110146 (2021).

11. Rathnayaka, Prabod & Mills, Nishan & Burnett, Donna & Silva, Daswin & Alahakoon, Damminda & Gray, Richard. A Mental Health Chatbot with Cognitive Skills for

Personalised Behavioural Activation and Remote Health Monitoring. Sensors. 22. 3653. 10.3390/s22103653 (2022).

12. Mrityunjay, Singh & Jakhar, Amit & Pandey, Shivam. Sentiment analysis on the impact of coronavirus in social life using the BERT model. Social Network Analysis and Mining. 11. 10.1007/s13278-021-00737-z (2021).

13. Nicol G, Wang R, Graham S, Dodd S, Garbutt J. Chatbot-Delivered Cognitive Behavioral Therapy in Adolescents with Depression and Anxiety During the COVID-19 Pandemic: Feasibility and Acceptability Study. JMIR Form Res. 2022 Nov 22; 6(11): e40242. doi: 10.2196/40242. PMID: 36413390; PMCID: PMC9683529.

14. Kapoor P, Agrawal P, Ahmad Z. Therapy chatbot: A relief from mental stress and problems. International Journal of Scientific & Engineering Research. 2021 Jul;12(5):1117.

15. DLA Piper Article. https://www.dlapiperdataprotection.com/?c=in, last accessed 2025/02/17.

16. Ministry of Electronics and Information Technology, Government of India. Draft Digital Personal Data Protection Rules, 2025. 2025 Jan 03, https://www.meity.gov.in/writereaddata/files/259889.pdf.

17. Timonin, Denis, BoYang Hsueh, and Vinh Nguyen. NVIDIA Article, https://developer.nvidia.com/blog/accelerated-inference-for-large-transformer-models-using-nvidia-fastertransformer-and-nvidia-triton-inference-server, last accessed 2022/08/03.

APPENDIX

1. Certificate for Conference of Research Paper at NIT Jalandhar

NITJ/ICE/E2ACON/2025/186

Springer

E2ACON 2025

International Conference on
Electrical, Electronics & Automation

Certificate of Participation

This is to certify that paper titled

Mental Health tracking and surveillance using CBT-based chatbot and BERT classifier

has been successfully presented by

Ananya Verma, KIET Group of Institutions, Ghaziabad, India

during International Conference on Electrical, Electronics and Automation held on 8-9 March 2025
organized by department of Instrumentation and Control Engineering,
Dr B R Ambedkar National Institute of Technology Jalandhar, Punjab

Dr. Afzal Sikander Organising Chair	Prof. Dilbag Singh General Chair	Prof. K S Nagla General Chair

2. Plagiarism Report

ORIGINALITY REPORT

20%	17%	12%	12%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	Submitted to KIET Group of Institutions, Ghaziabad Student Paper	3%
2	www.mdpi.com Internet Source	1%
3	formative.jmir.org Internet Source	1%
4	Submitted to HTM (Haridus- ja Teadusministeerium) Student Paper	1%
5	Submitted to University of Greenwich Student Paper	<1%
6	www.irejournals.com Internet Source	<1%
7	Submitted to University of North Texas Student Paper	<1%
8	www.hindawi.com Internet Source	<1%
9	ijisrt.com Internet Source	<1%
10	www.marcohealthtech.com Internet Source	<1%
11	www.ijitee.org Internet Source	<1%
12	Submitted to Harrisburg University of Science and Technology Student Paper	<1%