EMPATHI-AI – Conversational Chatbot for Mental Health Support

Prof S.S Kale[1], Kartik Bhosale[2], Sumit Bhosale[3], Pratik Dere[4], Tejas Aher[5,] Onkar Auti[6]
[1,2,3,4,5,6]*Department of Information Technology, JSCOE*

***Abstract*—Mental health support is a global need due to rising stress, anxiety and mental imbalance. We develop a chatbot providing empathetic mental health support using Agentic Artificial Intelligence (AI), Natural Language Processing (NLP) and Retrieval Augmented Generation (RAG) capabilities. Unlike symbolic chatbots, EMPATHI-AI has the capability to respond quickly with contextual and factual information based on expressive emotion detection and sentiment analysis, as well as the ability to find verified mental health information from reputable sources. The study stresses the system's potential to improve accessibility and privacy, as well as its contribution to emotional well-being and awareness through clever digital communication.**

***Index Terms*—Agentic AI, Mental Health Chatbot, Natural Language Processing (NLP), Emotion Detection, RAG, Digital Health.**

I. INTRODUCTION

Mental health has become a major global concern due to increasing stress, anxiety, and depression among individual. The majority of those suffering from mental illness are afraid to seek help because there is still substantial stigma attached to mental illness; the cost of treatment is too high; or there simply aren't enough therapists available. Therefore, there is a need for accessible digital solutions that can provide users with an initial level of emotional support.

Artificial Intelligence (AI) and Natural Language Processing (NLP) allow for the creation of conversational chatbots that can interact with users in the same way that a person would. As such, these chatbots can assist users through emotional guidance, stress management and mental health education.

The system proposed here, EMPATHI-AI, is an AI-based conversational chatbot designed to provide contextually-aided, supportive interactions. The EMPATHI-AI system encompasses the ability to detect user emotion, providing appropriate responses and will utilise NLP based models (chatbot) and Retrieval- Augmented Generation (RAG). All of these components are fundamental to provide reliable and empathetic responses in accordance with user privacy and ethical norms.

II. RESEARCH OBJECTIVES

The objectives of this study are:

- Build a conversational chatbot capable of holding discussions using NLP methods.
- Understanding user emotions through sentiment and emotion analysis
- Use Retrieval-Augmented Generation (RAG) to ensure accuracy.
- Provide compassionate support to users in context.
- To assess accuracy and overall participant acceptance of the system's responses

III. LITERATURE REVIEW

The rapid evolution of Artificial Intelligence (AI) Natural Language Processing (NLP) technology has also played a role in the emergence of many digital mental health systems that allow users to receive emotional support, early detection, or stress management through conversation-based technologies. In this section, we review the previous literature on which we base our proposed EMPATHI-AI system..

A. AI and Machine Learning in Mental Health:

Kasanneni et al. (2023) studied the problem of recognizing mental health disorders in social media

posts using machine and deep learning models. The models (BERT, RoBERTa, and XLNet) were trained to find language patterns that suggest the presence of anxiety, stress, and depression in the post text. Their findings showed high accuracy in classification; however, they were faced with computational constraints and could not depict multilingual or informal text as well. Nevertheless, this study shows that NLP could be a way to analyze mental health indicators; EMPATHI-AI is based on this premise.

B. Dialogue Systems for Mental Illness Detection:

Abilkaiyrkyzy et al. (2024) used the Rasa framework and fine-tuned BERT to create a prototype clever dialogue system for early detection of mental illness. They also used Artificial Intelligence and Clinical Psychology to classify the severity of depression and anxiety. The prototype was successful in a controlled setting but lacked emotion adaptability and natural language dialogue capabilities. This indicates the need for detecting emotion and adaptive natural language dialogue, which EMPATHI-AI attempts to provide through Agentic Artificial Intelligence.

C. Effective Analysis Of Machine and Deep Learning Methods for Diagnosing Mental Health using Social Media Conversations:

This includes the study of how machine and deep learning models can help ease and automate the detection of people who have mental illnesses by analyzing emotions, language, and sentiment in their online conversations using Natural Language Processing (NLP) techniques. State-of-the-art transformer models including BERT, RoBERTa, DistilBERT and XLNet were used to detect the signs of mental illness in the text extracted from social media posts. Results show that these state-of-the-art transformer models can learn and accurately detect mental illness, with several models attaining over 95% accuracy. This paper demonstrated that mining social media for public sentiment using advanced AI models helps effectively detect and implement timely mental health interventions.

D. Summary of Findings:

Recent evidence shows that AI is increasingly used to mediate mental health support but is limited in personalization, emotional understanding, and consistency. To address these limitations, we propose the EMPATHI-AI chatbot that uses emotion detection, agentic decision making, and retrieval-based
reasoning to conduct empathetic, context-aware, and trustworthy conversations.

IV. THEORETICAL FRAMEWORK

A. Foundation of EMPATHI-AI Framework

In the literature the theoretical framework EMPATHI- AI, which stands for Empathic AI based on Artificial Intelligence, NLP, Emotion Recognition, and HCI, was taken up and the aim of this framework was the development of a clever mental health chatbot that proposed to understand human emotions and respond with empathy using conversational AI.

In the theoretical foundation of mental health systems for checking verbal and non-verbal emotion, the system combines text-based sentiment analysis and facial emotion recognition to know users' emotions. In order to be able to implement a facial emotion recognition system, DeepFace is used. A deep neural network is used, trained with facial expression datasets, to detect emotions including happiness, sadness, anger, fear and neutrality. The system's emotions are used to adapt its behavior to the detected psychological state of the user.

B. Natural Language Processing and Emotion Detection

Along with visual emotion detection, the program includes Natural Language Processing (NLP) methods that analyze entered text using transformer models like BERT. It analyzes the structure of language, providing further understanding of emotional intent in the users' text input, the sentiment polarity, and any psychological distress the user appears to be experiencing. This dual emotion-detection mechanism increases the reliability of emotional interpretation when the camera or text input are individually used.

To generate its conversational intelligence, the chatbot employs a Large Language Model through the Gemini API. This allows it to generate contextually and emotionally appropriate responses, using past conversation history and emotional signals to give personalized or encouraging conversation. The

framework also includes response validation to avoid harmful or misleading content.

C. Risk Detection and Psychological Monitoring

Risk detection is a component of the model which seeks to detect people who are experiencing severe emotional distress, self-harm or suicidality. Tools can be created using natural language processing to detect high risk indicators in conversation. If such indicators are detected, the system brings up a safety protocol containing emergency support messages and a suggestion to contact professional mental health services or helplines.

The system also provides long-term psychological analysis by archiving the emotional content of interactions in a structured PostgreSQL database, periodically assessing mood patterns, and reporting on weekly mood trends to help identify chronic negative emotional patterns. This means the system is designed to serve as an early emotional support tool, not a medical diagnosis.

D. Security and Ethical Considerations

Security and privacy is another theoretical pillar of the framework. Secure login, password hashing, and JWT- based authentication secure user access to the system, which protects personal information such as chat history, emotion logs, and diary entries from access and misuse by others. Ultimately, the EMPATHI-AI system design framework for socially aware AI, conversational natural language processing (NLP), behavioral digital monitoring, and secure system design provides a supportive digital environment to assist an user in managing emotional stress and mental well-being.

It is important to note that the system is designed only to provide early emotional support or advice and that it cannot replace professional psychological or medical diagnosis.

V. METHODOLOGY

A. System Architecture and User Authentication

EMPATHI-AI is a layered architecture based on Artificial Intelligence (AI), Natural Language Processing (NLP), Emotion Recognition and secure data handling to develop a clever mental health support chatbot. It is focused on emotion detection, empathy generation, psychological trend monitoring, and secure communication to ensure privacy while offering emotional support to users.

User authentication and access to the system is the first stage. The user is directed to a secure authentication mechanism with password hashing and JWT based authentication to register and log into the system. This ensures that only authorized users can access the system and that sensitive data such as chat conversations, diaries, and emotional logs are protected.

B. Emotion Detection Mechanism

After users enter the system, they can interact with the chatbot by typing messages and sending an image via the camera of a connected device, with facial expression and text being the two types of emotional signals collected. The facial emotion detection mechanism is based on the DeepFace framework. DeepFace is a deep learning based facial recognition and analysis framework. The framework detects emotions such as happy, sad, angry, fearful and neutral. The system captures an image of the user's face using a webcam, and the DeepFace model processes the facial landmarks of the user's face to detect the emotion.

When the camera is unavailable or disabled, emotion inference is done using Natural Language Processing (NLP). A transformer-based model (such as BERT) analyzes messages sent by the user to infer sentiment polarity (positive/neutral/negative), intent (happiness/sadness/anger/etc.), and emotional distress signals. This dual-mode, multi-input emotion detection allows the system to continue detecting emotional state even when one input (camera or text) is unavailable.

C. Conversational Intelligence and Risk Detection

On identifying the user's sentiment, the detected sentiment and the user input is then sent to Gemini API that serves as the conversational intelligence engine of the platform. Using the provided sentiment data, the large language model is able to provide contextually aware responses that are attuned to the user's mood. It is intended to support an empathetic dialog while maintaining conversational continuity using stored chat history.

To avoid users having unsafe communication with the chatbot, the framework incorporates a response validation (output filtering) process, to be sure that

the chatbot does not output undesirable, incorrect, or inappropriate responses. The chatbot also has a disclaimer, notifying that the chatbot is not a licensed medical professional and is intended to provide emotional support.

Another important part of the system is risk detection. Using NLP for keyword detection as well as context analysis, the system can detect when an user is going through a crisis, is self-harming or has suicidal ideation. When such high-risk triggers are detected, we activate the safety protocol by showing emergency messages and directing the user to contact mental health professionals or helplines.

D. Data Storage, Psychological Tracking and Administration

All conversations are logged in a PostgreSQL database, in which a record is stored about users, chat logs, detected emotions, diary entries, and logs from risk assessment. This means the system can monitor for long-term patterns in an user's mood, enabling it to track an user's mood week-over-week and detect more persistent negative emotions to provide psychological support when needed. For psychological tracking and analytics, emotion logs are stored and long-term variations in an user's mood are detected to learn about their behavioral patterns and better interact with the user in future.

Finally, the system also includes an admin panel. Authorized administrators can monitor anonymized system analytics and risk alerts, as well as manage user accounts, without being able to view any sensitive user data.

Through emotion detection, conversational software, psychological tracking and security, EMPATHI-AI was proposed as a first step in emotional support in a secure, private and ethical manner. EMPATHI-AI is not intended to replace a psychologist or for medical diagnosis but rather be used as a first step in emotional support.

VI. CONCLUSION

EMPATHI-AI is an early emotional support chatbot based on AI, Natural Language Processing (NLP) and Emotion Recognition systems. It is based on DeepFace for facial emotion recognition, transformer-based NLP models such as BERT for text-based sentiment analysis, and the Gemini API for conversational artificial intelligence, and seeks to provide empathetic and context-aware replies. The system has resources to detect risk, including warnings for severe affect and suggesting an user seek additional professional input where necessary. The system requires authentication and ensures user privacy; it also tracks psychological trends. The overall goal of EMPATHI-AI is to be supportive and safe online, with explicit statements that use of the tool should never be a substitute for medical or psychological professional help, and should be seen as a first step in support.

VII. ACKNOWLEDGMENT

The authors would like to express their sincerest appreciation to Prof. S. S. Kale, Department of Information Technology, JSPM's Jayawantrao Sawant College of Engineering, Pune for her constant encouragement, expert guidance and valuable insights during the EMPATHI-AI project.

We would like to thank Prof. A. K. Gupta, Project Coordinator, Department of Information Technology, for her constant guidance, timely feedback and motivation for completing this work.

The authors wish to express their sincere gratitude to the Head of Department of Information Technology, Dr. S. R. Rangari, for his encouragement and valuable guidance, which inspired the authors to carry out the research successfully. The authors sincerely thank all the faculty and staff of Information Technology Department for their support constructive feedback, and technical assistance during the testing and evaluation phases.

REFERENCES

[1] Abilkaiyrkyzy, A, Laamarti, F., Hamdi, M., & ElSaddik, A. (2024). Dialogue System for Early Mental Illness Detection: Toward a Digital TwinSolution. IEEE Access, 12, 2007–2024. https://doi.org/10.1109/ACCESS.2023.3348783

[2] Hamdoun, S., Monteleone, R., Bookman, T.& Michael, K. (2023). AI-Based and Digital Mental Health Apps: Balancing Need and Risk. IEEE Technology and Society Magazine, 42(1), 25–36. https://doi.org/10.1109/MTS.2023.3241309

[3] Kasanneni, Y., Duggal, A., Sathyaraj, R., & Raja, S. P. (2025). Effective Analysis of Machine and

Deep Learning Methods for Diagnosing Mental Health Using Social Media Conversations. IEEE Transactions on Computational Social Systems,12(1),274-294.https://doi.org/10.1109/TCSS.2024.3487168

[4] Omarov, B., Narynov, S., & Zhumanov, Z. (2023). Artificial Intelligence-Enabled Chatbots in Mental Health: A Systematic Review. Computers, Materials & Continua, 74(3), 5105–5120. https://doi.org/10.32604/cmc.2023.034655

[5] Das, M., & Prasad, S. K. (2019). A Chatbot System for Mental Health Care. International Journal of Creative Research Thoughts (IJCRT), 7(2), 15–23. https://www.researchgate.net/publication/332845375_A_Chatbot_System_for_Mental_Health_C