iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 14 **Issue:** III **Month of publication:** March 2026

DOI: https://doi.org/10.22214/ijraset.2026.78368

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue III Mar 2026- Available at www.ijraset.com

MindCare - AI Powered Voice-Based Mental Health Companion

Prof. Nayana S. Thombare[1], Rohan Tirkhunde[2], Sneha Yamgar[3], Nikita Shinde[4], Vaishnavi Patole[5]
[1]*Lecturer,* [2, 3, 4, 5]*Student, Department of Computer Engineering, K J College of Engineering &Management Research Pune,India*

Abstract: ***Mental health challenges such as stress, anxiety, and depression are increasing worldwide, while access to timely and empathetic support remains limited due to social stigma, a shortage of professionals, and economic barriers. Recent advancements in artificial intelligence (AI) have enabled the development of digital mental health companions that provide scalable and accessible support. However, most existing systems rely heavily on text-based interaction, which may be inadequate during emotional distress when users struggle to type or clearly articulate their feelings. This research presents an AI-powered, voice-based mental health companion designed to provide empathetic and emotion-aware support through natural spoken interaction. The proposed system integrates speech-to-text processing, natural language processing (NLP), voice-based emotion recognition, and text-to-speech synthesis to detect users' emotional states and generate supportive, human-like responses. The system aims to enhance emotional expression, user engagement, and early mental health intervention. Experimental evaluation demonstrates that the proposed approach improves usability, emotional comfort, and interaction quality when compared to traditional text-based chatbots. The findings highlight the potential of voice-driven AI systems as effective complementary tools for promoting mental well-being.***

I. INTRODUCTION

Mental health has emerged as a critical global concern, affecting individuals across all age groups and social backgrounds. Factors such as academic pressure, workplace stress, social isolation, and rapid lifestyle changes have contributed to a significant increase in mental health issues, including anxiety, depression, and emotional burnout. Despite the growing demand for psychological support, many individuals hesitate to seek professional help due to social stigma, fear of judgment, financial limitations, and restricted access to mental health resources.With advancements in artificial intelligence, digital mental health solutions such as chatbots and virtual companions have emerged as alternative support systems. These tools offer advantages, including anonymity, continuous availability, and cost-effectiveness. However, most existing mental health chatbots rely primarily on text-based communication, which can be inconvenient and emotionally limiting, particularly during moments of distress, when users may find typing difficult or uncomfortable. Voice-based interaction provides a more natural and expressive mode of communication, allowing individuals to convey emotions through tone, pitch, and speech patterns. This research focuses on developing an AI-powered voice-based mental health companion capable of understanding emotional cues from speech and responding empathetically. By integrating speech recognition, emotion detection, and conversational artificial intelligence, the system aims to bridge the gap between human-like emotional support and digital accessibility

II. LITERATURE REVIEW

Numerous AI-driven mental health chatbots have been developed to assist users with emotional regulation, stress management, and self-reflection. Popular systems such as Woebot, Wysa, and Replika utilize natural language processing (NLP) and predefined therapeutic frameworks, including Cognitive Behavioral Therapy (CBT), to guide conversations. These systems have demonstrated effectiveness in improving short-term emotional awareness and coping strategies. However, their dependence on text-based interaction limits emotional depth and reduces usability during periods of intense distress. Voice emotion recognition has emerged as a critical research area in affective computing. Emotional states can be identified by analyzing acoustic features such as Mel-Frequency Cepstral Coefficients (MFCCs), pitch, energy, spectral centroid, and speech tempo. Machine learning and deep learning approaches, including Support Vector Machines (SVM), Convolutional Neural Networks (CNNs), and Long Short-Term Memory (LSTM) networks, have demonstrated high accuracy in classifying emotions such as happiness, sadness, anger, fear, and neutrality. Despite promising results, many studies primarily focus on emotion classification without integrating conversational intelligence and real-time user interaction.

©IJRASET: All Rights are Reserved

Existing research highlights the effectiveness of AI-based chatbots and emotion recognition techniques independently; however, there is a lack of integrated systems that combine voice-based emotion detection with empathetic conversational response generation. Furthermore, aspects such as personalization, emotional continuity, and real-time voice interaction remain underexplored.

These limitations motivate the development of a unified, voice-driven, emotion-aware mental health companion. The proposed system, MindCare- AI-powered voice-based mental health companion, is designed as a modular, scalable, and real-time architecture that enables natural voice interaction, emotional intelligence, and empathetic conversational support. The architecture integrates speech processing, emotion recognition, natural language understanding, and response generation to deliver seamless human-like interaction. Fig. 1 illustrates the overall system architecture.

III. SYSTEM OVERVIEW

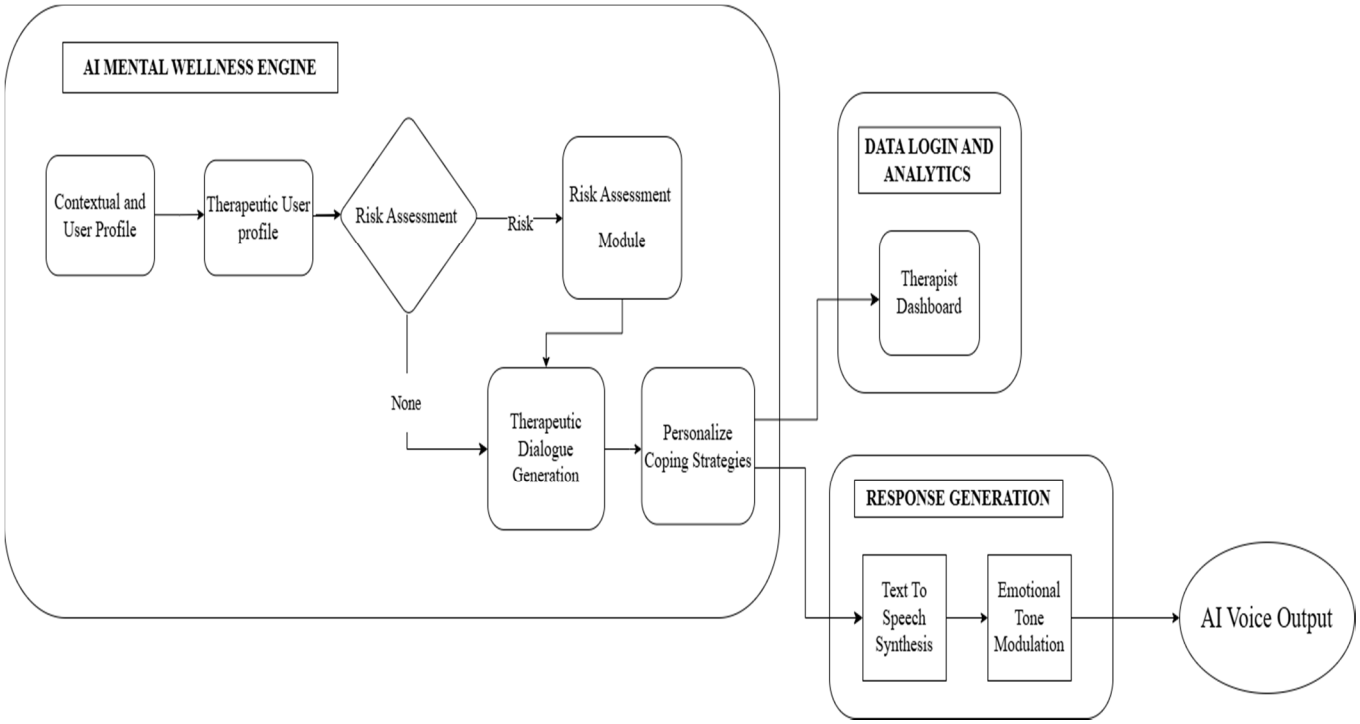
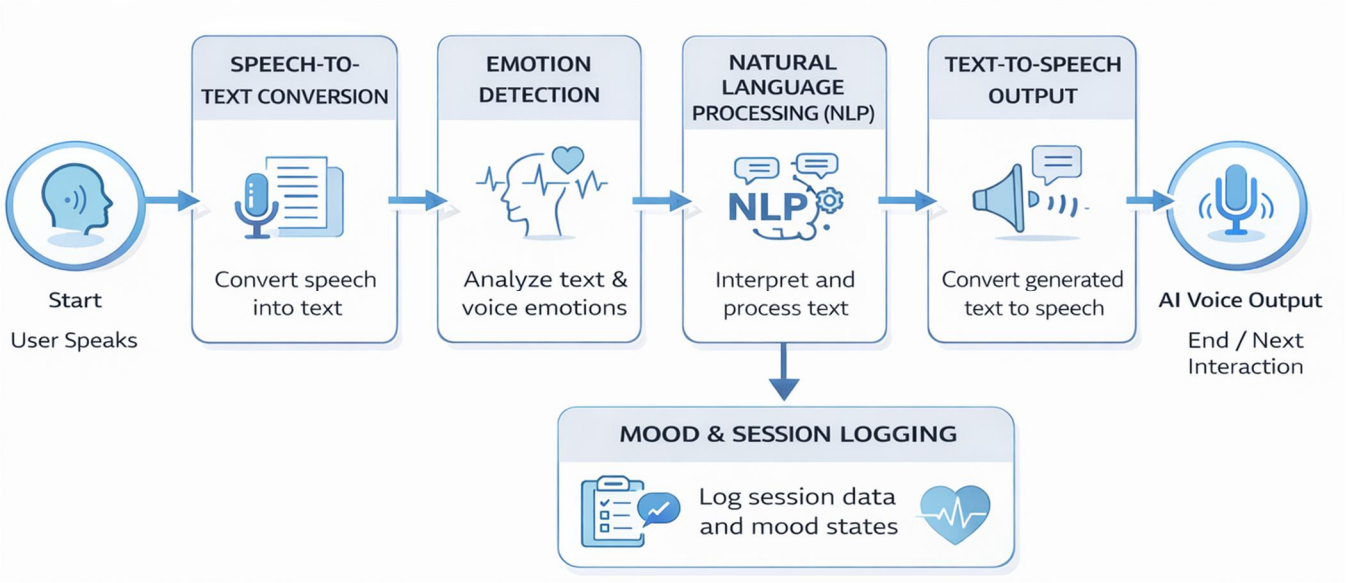


Fig1. Block Diagram

The proposed system architecture presents an AI-driven mental wellness engine designed to provide personalized psychological support through intelligent dialogue and voice interaction. The process begins with the Contextual and User Profile module, where user information and contextual data are collected to understand the individual's emotional and behavioral state. This data is used to construct a Therapeutic User Profile, which represents the user's psychological characteristics and interaction history. Next, the system performs Risk Assessment to determine whether the user exhibits signs of psychological distress. If potential risk is detected, the Risk Assessment Module conducts a deeper evaluation to determine the severity of the condition. If no critical risk is identified, the system proceeds to Therapeutic Dialogue Generation, where supportive and empathetic responses are generated using natural language processing techniques. These responses are further enhanced through Personalized Coping Strategies, which provide tailored recommendations to help users manage stress and emotional challenges. The system also includes a Data Logging and Analytics module that records user interactions and emotional indicators. This information can be accessed through a Therapist Dashboard to support monitoring and professional intervention when required. Finally, the Response Generation module converts the generated text into speech using Text-to-Speech synthesis and adjusts the delivery using Emotional Tone Modulation to ensure empathetic communication. The response is then delivered to the user as AI Voice Output, enabling a natural and interactive support experience.

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue III Mar 2026- Available at www.ijraset.com

IV. METHODOLOGY



This section presents the methodology adopted for developing MindCare— AI-powered voice-based mental health companion. The system integrates speech processing, emotion recognition, natural language understanding, and intelligent response generation to deliver real-time, emotionally adaptive, and secure user interaction.

A. System Workflow

The system follows a closed-loop conversational framework consisting of voice capture, audio preprocessing, speech-to-text conversion, emotion detection, natural language processing, response generation, and speech synthesis. This modular architecture ensures accurate understanding, low latency, and empathetic interaction.

B. Voice Processing and Emotion Recognition

User speech is captured through device microphones and enhanced using noise reduction and signal normalization techniques. The refined audio is transcribed using deep learning-based speech recognition models. Simultaneously, emotional features such as MFCCs, pitch, and energy levels are extracted and classified to determine the user's emotional state.

C. Natural Language Understanding and Context Analysis

The transcribed text undergoes semantic and contextual analysis using NLP techniques to identify intent and sentiment. Dialogue context and interaction history are maintained to ensure conversational continuity and personalized support.

D. Intelligent Response Generation and Speech Synthesis

Emotion-aware responses are generated using transformer-based conversational models and converted into expressive speech through neural text-to-speech synthesis, enabling natural and empathetic communication.

E. Data Security and Privacy

Strong encryption, anonymization, and access control mechanisms are implemented to protect sensitive user data and ensure ethical system operation.

V. FUTURE SCOPE

The proposed system can be further enhanced by incorporating multilingual voice interaction to support a wider user base. Advanced emotion prediction models based on temporal learning can improve emotional understanding over extended conversations. Personalization through adaptive learning can enable tailored emotional support for individual users. Additionally, integration with wearable devices and collaboration with mental health professionals can enhance real-time monitoring and clinical relevance.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

VI. CONCLUSION

This research presented MindCare- AI-powered voice-based mental health companion designed to deliver empathetic and emotionally intelligent support through natural voice interaction. By integrating speech processing, emotion recognition, and natural language understanding, the system enables real-time, personalized, and context-aware emotional assistance. Experimental observations indicate improved emotional expressiveness, user comfort, and engagement, demonstrating the system's potential as an effective tool for early-stage mental health support and emotional well-being.

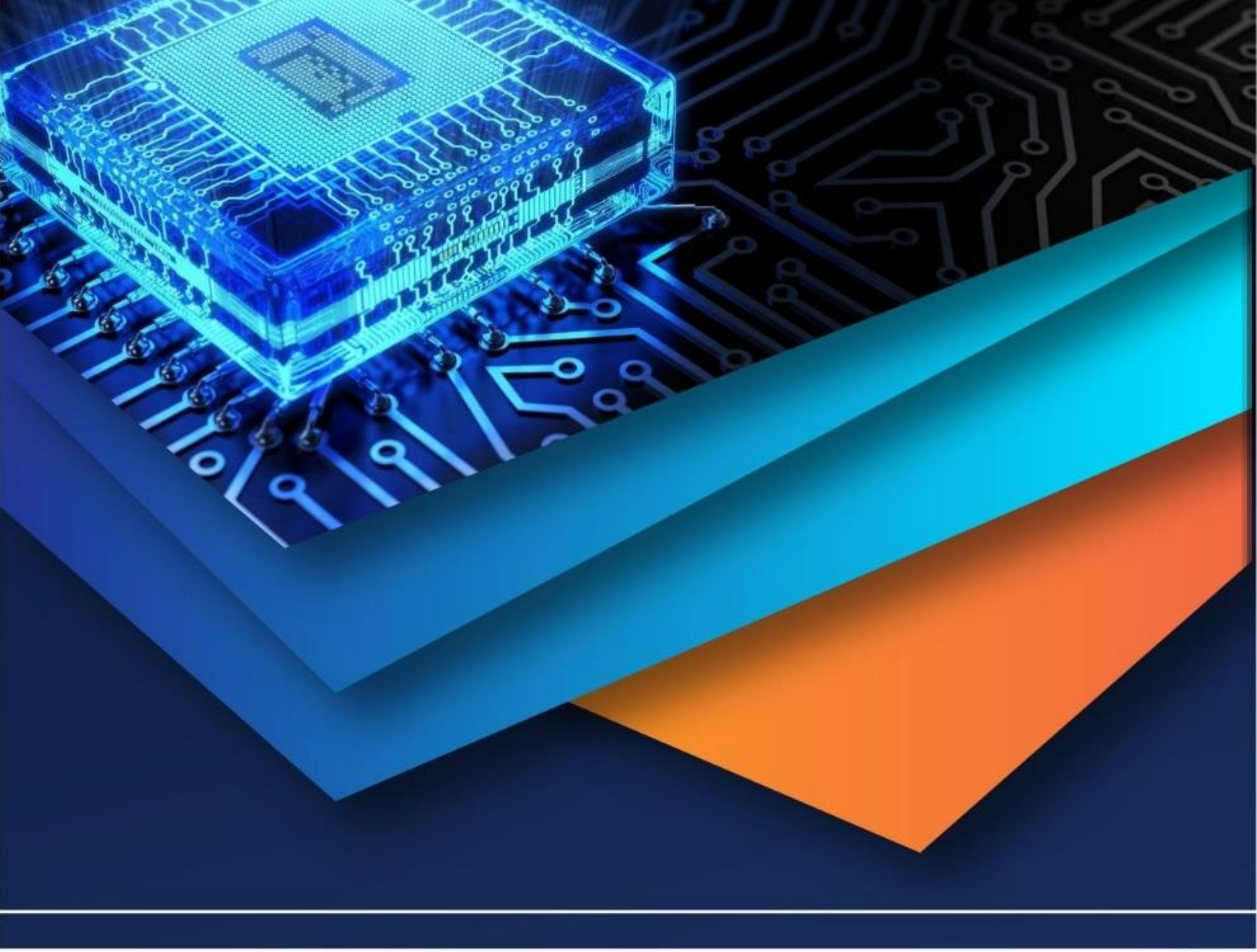
VII. ACKNOWLEDGMENT

The authors gratefully acknowledge the guidance of Prof. Nayana S. Thombare and the Department of Computer Engineering, K. J. College of Engineering and Management Research, Pune, for their invaluable support.

REFERENCES

[1] World Health Organization, "Mental health action plan 2013–2030," WHO Press, Geneva, Switzerland, 2021.
[2] S. Fitzpatrick, A. Darcy, and M. Vierhile, "Delivering cognitive behavior therapy to young adults with symptoms of depression and anxiety using a fully automated conversational agent (Woebot): A randomized controlled trial," JMIR Mental Health, vol. 4, no. 2, pp. 1–10, 2017.
[3] A. Miner, L. Milstein, S. Schueller, R. Hegde, C. Mangurian, and E. Linos, "Smartphone-based conversational agents and responses to user expressions of emotional distress," JAMA Internal Medicine, vol. 176, no. 5, pp. 619–625, 2016.
[4] B. Schuller, S. Steidl, and A. Batliner, "The Interspeech 2013 computational paralinguistics challenge: Social signals, conflict, emotion, autism," in Proc. Interspeech, Lyon, France, 2013, pp. 148–152.
[5] Z. Zhang, J. Han, E. Coutinho, and B. Schuller, "Dynamic difficulty awareness training for continuous emotion prediction," IEEE Transactions on Multimedia, vol. 21, no. 5, pp. 1280–1293, 2019.
[6] T. Young, D. Hazarika, S. Poria, and E. Cambria, "Recent trends in deep learning-based natural language processing," IEEE Computational Intelligence Magazine, vol. 13, no. 3, pp. 55–75, 2018.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)