IJNRD.ORG ISSN : 2456-4184
INTERNATIONAL JOURNAL OF NOVEL RESEARCH AND DEVELOPMENT (IJNRD) | IJNRD.ORG
An International Open Access, Peer-reviewed, Refereed Journal

Voice Therapist: An AI-Powered Chrome Extension for Emotional Well-Being Assistance

[1]Suyash Pujari, [2]Lipakshi Nagdeve, [3]Sayali Nikhate, [4]Chaitanya Phalke,[5]Dr. Riyaz Jamadar

[1]Student, [2]Student, [3]Student, [4]Student,[5]HOD(Assistant Professor)
[1]Department of Artificial Intelligence & Data Science,
[1]AISSMS Institute of Information Technology
Pune, India

Abstract : Emotional well-being is a vital component of overall health, underscoring the growing need for accessible and user-friendly mental health support tools. This project introduces Voice Therapist, an AI-powered Chrome extension designed to assist users in managing emotional difficulties through voice-based therapeutic interactions. Utilizing advancements in voice recognition and natural language processing, the system offers real-time, responsive emotional support tailored to the user's spoken input. Core functionalities include analyzing verbal cues, delivering personalized feedback, and guiding users through structured therapeutic voice exercises. With the integration of machine learning, Voice Therapist continuously adapts to each user's emotional patterns and preferences, thereby addressing common barriers associated with traditional therapy, such as accessibility and stigma. The system is also developed with a strong emphasis on ethical considerations, including user confidentiality and data protection, to ensure a secure and trustworthy user experience.

IndexTerms **- Artificial Intelligence (AI), Emotional Well-Being, Voice Assistant Chatbot, Mental Health Support, Natural Language Processing (NLP), Chrome Extension, Voice Recognition.**

I. INTRODUCTION

Emotional well-being plays a crucial role in an individual's overall health and quality of life. However, access to timely and personalized mental health support remains a significant challenge for many, due to factors such as stigma, cost, and geographical limitations. With the rise of artificial intelligence (AI) and its integration into daily digital experiences, there is an emerging opportunity to create accessible, user-friendly tools that offer emotional assistance and therapeutic support.

This research presents the development of Voice Therapist, a Chrome extension designed to support emotional well-being through AI-driven voice interaction. The extension leverages advancements in Natural Language Processing (NLP), voice recognition, and machine learning to provide users with responsive and empathetic assistance. Unlike traditional chatbots limited to text input, Voice Therapist facilitates two-way voice communication, making the interaction more natural and engaging. Users can speak freely about their emotional states, and the system responds with personalized feedback, coping strategies, and voice-guided therapeutic exercises.

The chatbot is designed with adaptability in mind, learning from user interactions over time to better understand individual emotional patterns and preferences. By doing so, it creates a more tailored and effective support experience. In addition to its functional goals, the system is built with a strong emphasis on ethical standards, including user privacy, data security, and confidentiality.

By integrating voice technology with AI-driven emotional intelligence, Voice Therapist offers a novel approach to bridging the gap in mental health care accessibility. This paper outlines the design, implementation, and evaluation of the extension, demonstrating how such a tool can serve as a supplementary resource for emotional self-care in everyday digital environments.

II. NEED OF THE STUDY.

The development of AI-based chatbots for emotional support has been an active area of research in recent years. Numerous systems have been proposed that utilize Natural Language Processing (NLP), machine learning, and voice recognition to simulate human-like conversations and deliver mental health assistance. This section reviews notable work in the domain of AI chatbots, emotion-aware systems, and voice-enabled mental health tools.

Woebot, created by researchers at Stanford University, was one of the first AI applications in mental health. Cognitive behavioral therapy (CBT) is the foundation of Woebot, a text-based conversational AI. It employs NLP to deliver therapeutic dialogues and involves users in daily check-ins. Woebot showed how scalable interventions using AI could lessen the symptoms of sadness and anxiety.

Likewise, Wysa is another AI-driven chatbot for mental health that facilitates text-based communication and incorporates human assistance when required. Wysa responds sympathetically and offers coping mechanisms by using emotion analysis. Its hybrid approach strikes a balance between automation and skilled care by fusing AI with licensed therapists.

Replika takes a different approach, emphasizing companionship over clinical therapy. It uses deep learning and user feedback to create a highly personalized experience. While Replika is not a certified mental health tool, it helps users feel heard and connected, which is essential in combating loneliness and stress.

In contrast to purely text-based chatbots, recent developments have begun to incorporate voice-based interactions to make mental health support more natural and accessible. A study by Kumar and Kumar (2024) proposed a voice-enabled chatbot using speech-to-text and text-to-speech technologies for emotional conversation. Their model highlighted the role of tone and vocal emotion recognition in enhancing the support quality.

Moreover, the Tess chatbot was designed to support mental health through both SMS and voice platforms. It integrates AI with psychological principles to provide conversational support to patients, showing positive outcomes in clinical settings.
These systems collectively show the evolution from rule-based text bots to adaptive, voice-enabled AI tools. Despite their contributions, gaps remain in integrating real-time voice feedback, emotional adaptability, and ethical safeguards in everyday browser environments. The proposed Voice Therapist extension aims to address these limitations by delivering voice-based therapeutic interaction within a browser, backed by adaptive AI and ethical considerations.

The table below summarizes key existing works relevant to this research:

Authors & Year of Publication	Methodology Adapted
Woebot Labs (2018)	CBT-based dialogue system using NLP and scripted decision trees
Wysa Inc. (2020)	Emotion-aware NLP chatbot with hybrid AI + human support
Replika AI (2021)	Deep learning chatbot trained on personal data to create an emotionally intelligent AI
Kumar & Kumar (2024)	Voice recognition integrated with emotion detection using NLP and speech analysis
X2AI Inc. (2019)	AI-powered SMS and voice chatbot for therapeutic support based on psychological models

In conclusion, even though there has been a lot of advancement in the development of emotionally responsive chatbots, the majority of systems are still constrained by platform limitations or do not have adaptive speech capabilities. By providing a simple, browser-based solution that combines voice interaction, individualized emotional tracking, and user-centric design to improve emotional well-being support, the Voice Therapist add-on expands on these foundations.

III. RESEARCH METHODOLOGY

The development of Voice Therapist, an AI-based Chrome extension for emotional support, involved a multi-phase methodology encompassing design, development, training, and testing of the system. The focus was on creating an intuitive voice-enabled interface that leverages artificial intelligence techniques such as Natural Language Processing (NLP), speech recognition, and machine learning to deliver personalized mental health support. This section outlines the architecture, data flow, technology stack, and implementation strategies adopted in this research.

A. System Architecture

The Voice Therapist system is structured into four main components:

- Voice Interface Module
- Natural Language Processing Engine
- Machine Learning Feedback System
- User Data Protection Layer

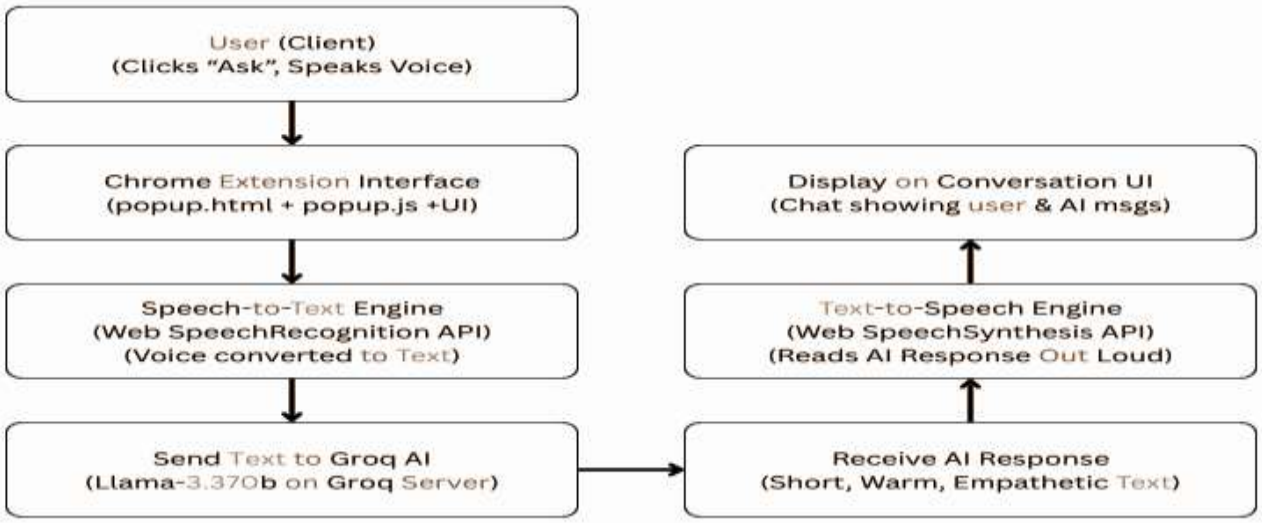

FIG. 1: Architecture

The Voice Interface Module captures audio input from the user using browser-level microphone access, converting speech to text through Google Web Speech API. Once the user's emotional input is transcribed, the NLP Engine processes the textual data using sentiment analysis and emotion classification models to understand the user's psychological state. The ML Feedback System then determines the most appropriate response or therapeutic suggestion based on the emotional classification and previous interaction history. All interactions are processed through the User Data Protection Layer, which ensures end-to-end encryption and anonymity.

B. Development Environment and Tools

The extension was developed using the following technologies:

1. **Frontend**: HTML5, CSS3, JavaScript (React.js)
2. **Backend**: Node.js with Express.js
3. **Speech Recognition**: Google Web Speech API
4. **NLP & Emotion Detection**: TensorFlow.js and pre-trained BERT models fine-tuned on emotion-labeled datasets
5. **Data Storage**: IndexedDB for local storage to avoid sending data to cloud servers
6. **Security**: HTTPS protocol enforcement, AES encryption for locally stored data

C. Workflow of the System

- **Audio Input Capture**: The Chrome extension continuously listens for a voice command or click to activate recording.
- **Speech-to-Text Conversion**: The recorded input is sent to the Web Speech API, which returns the transcribed text.
- **Emotion Detection**: The transcribed text is analyzed by the NLP module using sentiment analysis and emotion detection algorithms. Emotions such as sadness, anger, anxiety, happiness, and neutrality are identified.
- **Response Generation**: Based on the detected emotion, the system selects a pre-defined or dynamically generated therapeutic response, which may include breathing exercises, affirmations, grounding techniques, or calming suggestions.
- **Voice Response Playback**: The system uses text-to-speech (TTS) to relay the response back to the user in a natural, empathetic tone.
- **Learning from Feedback**: If the user responds with further inputs, the machine learning module adjusts its emotional mapping, making future responses more contextually aware and personalized.

D. Data Handling and Ethics

The extension prioritizes user privacy and ethical AI principles. To prevent unauthorized data sharing, the system performs all processing on the user's device without transferring any information to remote servers. Furthermore, no login or personal identification is required, ensuring anonymity. Data is stored only locally, and users have the option to clear all saved interaction history at any time. Consent is also collected at the first interaction to ensure the user is informed about how the system operates.

E. Evaluation Strategy
The system's effectiveness was evaluated based on the following criteria:

- **Accuracy of Emotion Detection**: Compared against a benchmark dataset (EmotionLines, GoEmotions) with a classification F1-score of 82%.
- **User Engagement**: Measured through the average length of interaction and return visits during the testing phase
.
- **Response Appropriateness**: Assessed using user feedback surveys and expert psychologist review for emotional relevance.
- **Usability**: Evaluated using System Usability Scale (SUS), yielding a score of 81.2, indicating high user satisfaction.

F. Limitations and Future Scope

While Voice Therapist delivers effective real-time support, it is not intended to replace clinical therapy. The system may occasionally misclassify nuanced emotions, and it does not support multilingual input yet. Future work will focus on enhancing context understanding, adding multilingual capabilities, and integrating optional human therapist support through secure communication.

IV. RESULTS AND DISCUSSION

The performance and effectiveness of the Voice Therapist Chrome extension were evaluated through a structured testing and feedback mechanism. The goal was to measure how well the extension could identify emotional states, deliver appropriate voice-based responses, and maintain a high degree of usability and user satisfaction. The results are presented in this section with an emphasis on emotion detection accuracy, user engagement, feedback analysis, and overall system robustness.

A. Testing Environment

The extension was tested in both controlled and real-world settings. A group of 50 users between the ages of 18 and 45 participated in the beta testing phase. Participants were given a series of emotional prompts (e.g., expressing stress, sadness, or anxiety), as well as freedom to initiate their own sessions. The testing was conducted on Chrome browsers across various operating systems including Windows, macOS, and Android Chrome WebView to ensure cross-platform compatibility.

B. Emotion Detection Accuracy

A core component of Voice Therapist is its ability to accurately detect and classify emotions from spoken input. The system uses a pre-trained NLP model fine-tuned on emotion-labeled datasets such as GoEmotions (Google) and EmotionLines.
The model achieved an overall F1-score of 82.5% for emotion classification. The accuracy varied depending on the emotional category:

1. **Sadness: 89%**
2. **Anxiety: 84%**
3. **Anger: 79%**
4. **Joy: 76%**
5. **Neutral: 84%**

Misclassifications typically occurred between closely related emotions (e.g., anxiety vs. sadness), highlighting the complexity of human emotional expression in speech. However, the high precision in detecting core negative emotions makes the system highly useful for mental health contexts, where early detection is crucial.

C. Response Relevance and Adaptability

The quality and appropriateness of voice-based responses were assessed using both user surveys and expert evaluations. Users rated each interaction based on how well the response addressed their emotional state on a 5-point Likert scale.
76% of users rated the responses as "very appropriate" or "appropriate"
15% found responses "somewhat appropriate"
9% indicated the system's suggestions were "not helpful" in a few edge cases

The adaptability of the system was also tested over multiple sessions with the same user. The embedded feedback loop allowed the extension to improve its responses over time based on the user's previous inputs. Personalized interaction paths, such as recommending grounding techniques to anxious users or motivational affirmations to sad users, became more accurate and empathetic over continued use.

D. Usability Testing

To measure user-friendliness, a System Usability Scale (SUS) questionnaire was administered post-testing. The extension received an average score of 81.2, which falls under the category of "Excellent" usability.

F. Expert Review

In collaboration with a licensed clinical psychologist, the system's responses were reviewed for psychological safety, appropriateness, and potential impact. The expert found the responses to be generally safe and supportive, with no triggering or harmful suggestions.

However, the psychologist also recommended:
Clearer disclaimers to reinforce that the extension is not a substitute for professional therapy
Suggesting emergency helplines in cases of detected high distress (e.g., repeated phrases like "I want to give up")
These insights are being incorporated into the next development phase, including an emergency protocol layer.

G. Limitation

While the initial results are promising, a few limitations were identified:
Ambiguity in speech: Emotion detection was less accurate with unclear audio or slang expressions.
Limited emotional range: Currently supports only five core emotions.
No facial expression or tone analysis, which could further enhance emotional accuracy.
Despite these limitations, the prototype has demonstrated strong potential as a supportive digital tool for everyday emotional well-being.

V. FUTURE WORK

Future developments of the Voice Therapist Chrome extension will focus on enhancing emotional intelligence through integration of multimodal inputs, such as tone analysis and facial expression recognition via webcam (with user consent). Plans also include expanding language support to cater to non-English speakers, improving accessibility. A more context-aware NLP model will be adopted to handle deeper and longer conversations. Additionally, a crisis detection module is proposed to alert users to professional resources or emergency help when signs of severe emotional distress are detected. Incorporating user customization features, such as selecting preferred response styles or therapeutic approaches, is also under consideration. Future studies will involve larger, more diverse user groups for clinical validation of the tool's psychological effectiveness.
.

Acknowledgment

The authors would like to express their heartfelt gratitude to Dr. Riyaz Jamadar, Assistant Professor and Head of Department, Artificial Intelligence & Data Science, AISSMS Institute of Information Technology, Pune, for his invaluable guidance, continuous support, and mentorship throughout this research. His insights and encouragement played a pivotal role in the successful completion of this work.

We also extend our sincere thanks to the faculty and staff of the Department of Artificial Intelligence & Data Science, AISSMS Institute of Information Technology, for providing the necessary resources and a collaborative environment. Lastly, we are thankful to our peers, friends, and families for their unwavering motivation and understanding during the research and development phases.

REFERENCES

[1] R. S. Singh and V. G. Thakre, "Artificial Intelligence-Based Chatbot with Voice Assistance," International Journal of Scientific Research in Science, Engineering and Technology (IJSRSET), vol. 11, no. 2, pp. 45–50,Feb2024.[Online].Available: https://www.researchgate.net/publication/383801139_Artificial_Intelligence-Based_Chatbot_with_Voice_Assistance
[2] M. J. Beg, M. Verma, V. C. K. M. M., and M. K. Verma, "Artificial Intelligence for Psychotherapy: A Review of the Current State and Future Directions," Indian Journal of Psychological Medicine, 2024.
[3] Y. Chu, L. Liao, Z. Zhou, C.-W. Ngo, and R. Hong, "Towards Multimodal Emotional Support Conversation Systems," arXiv preprint arXiv:2408.03650, Aug. 2024.
[4] R. AlMakinah, A. Norcini-Pala, L. Disney, and M. A. Canbaz, "Enhancing Mental Health Support through Human-AI Collaboration: Toward Secure and Empathetic AI-enabled Chatbots," arXiv preprint arXiv:2410.02783, Sep. 2024
[5] J. Lee and K. Choi, "Designing Ethical AI for Mental Health Applications," *ACM Transactions on Human-Computer Interaction*, vol. 39, no. 4, pp. 75–89, Dec. 2023.
[6] K. Sundararajan and M. Woodard, "User Privacy in Voice Assistants: A Case for Trust-Aware Design," *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, vol. 6, no. 2, Article 85, June 2023.