

EMPATH.AI: A CONTEXT-AWARE CHATBOT FOR EMOTIONAL DETECTION AND SUPPORT

1st Neave Kallivalappil
Computer Engineering Department
Fr. Conceicao Rodrigues
College of Engineering
Mumbai, India
neave.mailbox@gmail.com

2nd Kyle D'souza
Computer Engineering Department
Fr. Conceicao Rodrigues
College of Engineering
Mumbai, India
kyle.dsouza.official@gmail.com

3rd Afif Deshmukh
Computer Engineering Department
Fr. Conceicao Rodrigues
College of Engineering
Mumbai, India
afifkhanab@gmail.com

4th Chinmay Kadam
Computer Engineering Department
Fr. Conceicao Rodrigues
College of Engineering
Mumbai, India
chinmaykad172@gmail.com

5th Neha Sharma
Data and Analytics Practice
Tata Consultancy Services
Pune, India
nvsharma@rediffmail.com

Abstract—The increasing prevalence of mental health disorders highlights the urgent need for innovative technologies that can provide support and improve individuals' emotional well-being. Intelligent Virtual Friends (IVFs) have emerged as a promising solution to offer personalized and effective therapeutic experiences. This study introduces Empath.ai, an emotionally intelligent multimodal chatbot therapist powered by machine learning. Empath.ai aims to recognize and respond to users' emotions, providing a unique and tailored experience by combining cognitive-behavioural therapy techniques with empathetic communication. This integration of facial, text, and voice emotion recognition empowers the chatbot to deliver a holistic and tailored approach to emotional support, ensuring that users receive the assistance they need in a comprehensive and empathetic manner and serves as a valuable resource for those seeking accessible alternatives to traditional therapy methods.

Index Terms—Mental health, Chatbot therapist, Cognitive-behavioral therapy, Artificial intelligence, Empathetic communication

I. INTRODUCTION

A. Background

Mental health issues are a growing concern worldwide, affecting millions of individuals. Studies indicate that approximately 1 in 5 people experience a mental or neurological disorder at some point in their lives, with depression being a leading cause of disability globally [1]. The COVID-19 pandemic has further exacerbated the situation, leading to increased stress, anxiety, and depression due to factors such as isolation, job loss, and financial insecurity [2].

Despite the availability of effective treatments, many individuals do not seek or receive appropriate care due to various reasons, including limited resources, stigma, and other

barriers. Consequently, there exists a significant treatment gap, with approximately 60% of people with mental disorders not receiving any form of treatment [3]. Moreover, even when individuals do seek treatment, they may encounter long waiting lists for therapy or face limited access to qualified mental health professionals [4].

Empath.ai stands as an innovative multimodal chatbot that goes beyond traditional text-based systems. It harnesses a range of emotion recognition technologies to perceive and interpret the emotional states exhibited by its users. Through the integration of facial emotion recognition, text-based emotion extraction, and speech emotion recognition, this chatbot adeptly identifies and addresses the emotional needs of each individual.

B. Motivation

In response to these challenges, there has been an increasing interest in utilising technology to provide mental health support and therapy. One promising approach is the use of chatbots, which have demonstrated effectiveness in offering emotional support and therapy to individuals with mental health conditions [5]. Chatbots provide a convenient and accessible avenue for individuals to seek help at any time and from any location, without the need for appointments or physical presence.

The chatbot can incorporate various therapeutic techniques, such as cognitive-behavioural therapy, mindfulness, and positive psychology, to assist users in managing their symptoms.

What sets Empath.ai apart is its context-awareness, which enables it to tailor its responses to the user's specific situation and environment. The chatbot can perceive the user's mood,

tone, and sentiment, enabling it to adapt its replies accordingly. For instance, if a user expresses feelings of anxiety, Empath.ai can offer relaxation exercises or coping strategies. Similarly, if a user reports feeling overwhelmed, the chatbot can suggest self-care activities or encourage taking a break.

Empath.ai has the potential to bridge the treatment gap and provide accessible and effective mental health support to millions of individuals worldwide. Given the urgency highlighted by the COVID-19 pandemic to address mental health challenges, Empath.ai's context-aware approach offers a promising solution for the future.

II. LITERATURE REVIEW

This section situates the study within the larger body of literature on chatbots for mental health support. It provides an overview of relevant studies on the design, implementation, and evaluation of chatbots in the context of mental health care. A comprehensive search was performed using relevant keywords and databases. The inclusion criteria focused on studies that examined the use of chatbots for mental health support, specifically in relation to emotional well-being, symptom monitoring, and cognitive-behavioral therapy. The literature on chatbots for mental health reveals a growing body of research in this field. Several studies have investigated the design aspects of chatbots, including their conversational agents framework and their ability to recognize and respond to emotions [5], [6].

Evaluation studies have focused on assessing the effectiveness of chatbots in improving mental health outcomes. These studies have shown promising results, with chatbots demonstrating efficacy in reducing symptoms of depression and anxiety [5]. User perceptions of chatbots have also been explored, highlighting the importance of empathetic communication and context-awareness in enhancing user satisfaction and engagement [7], [8].

Relevant Literature related to Design, Implementation and Evaluation of Chatbots:

Design of Chatbots for Mental Health: Chatbot design involves considering the target audience, specific mental health conditions, and therapeutic techniques. Incorporating conversational agents equipped with emotion recognition capabilities and context-awareness has been shown to enhance user engagement and satisfaction [6].

Implementation of Chatbots for Mental Health: Effective implementation of chatbots for mental health requires the integration of technologies such as natural language processing, machine learning, and emotion recognition. Emotion recognition technology enables chatbots to detect and interpret users' emotional states, leading to tailored responses and interventions. Furthermore, the use of sentiment analysis techniques can aid in understanding users' emotional well-being and delivering personalized support [9].

Evaluation of Chatbots for Mental Health: Numerous studies have assessed the effectiveness of chatbots for mental health, examining their impact on user outcomes such as

symptom alleviation and well-being improvement. For instance, research on chatbots designed for depression using cognitive-behavioral therapy techniques has shown significant improvements in symptoms [5]. User perceptions of empathy and engagement have also been found to be crucial factors contributing to the success of mental health chatbots [6], [8].

The existing body of work on chatbots for mental health serves as a valuable foundation for the development and evaluation of Empath.ai, a context-aware chatbot designed to deliver emotional support and therapy. By leveraging the insights gained from previous research, Empath.ai aims to bridge the gap in mental health care and provide individuals with a personalised and effective means of managing their emotional well-being.

Based on the relationship between these themes and the study of Empath.ai, the following hypotheses are proposed:

Hypothesis 1: The integration of emotion recognition and context-awareness in Empath.ai will lead to increased user engagement and satisfaction compared to traditional chatbots.

Hypothesis 2: Users of Empath.ai will experience greater improvements in mental health outcomes, including symptom alleviation and well-being enhancement, compared to users of traditional chatbots.

Chatbots hold immense promise in addressing the treatment gap in mental health and providing accessible and efficient mental healthcare services. Incorporating emotion recognition technology and context-awareness has proven to be instrumental in enhancing the effectiveness of mental health chatbots. Moreover, user perceptions of empathy and engagement play a vital role in their overall success.

The primary objective of this research is to assess the effectiveness of Empath.ai in delivering emotional support and therapy to users. Specifically, we aim to investigate whether Empath.ai's capacity to identify and respond to emotional states results in improved mental health outcomes when compared to traditional chatbots. Additionally, we seek to explore user perceptions of Empath.ai, including their level of trust in the chatbot, satisfaction with the interaction, and willingness to utilise it again in the future. By achieving these objectives, our aim is to contribute to the advancement of more effective and accessible mental health care solutions utilising chatbot technology.

III. MATERIALS AND METHODS

A. Data Collection

The development of Empath.ai required a comprehensive dataset that covered various aspects of mental health. To achieve this, a meticulous data collection process was carried out, focusing on conversations related to mental health from diverse outlets. The dataset comprised discussions addressing a wide range of mental health concerns such as anxiety, depression and many others.

The collected dataset was utilised to train and fine-tune the different components of Empath.ai, including the natural language processing (NLP), facial emotion recognition, and audio-based emotion recognition models. The NLP algorithm

learned patterns and nuances in mental health conversations, enabling the chatbot to understand user queries and sentiments.

In addition to the text-based conversations, the collected datasets also included facial expressions and audio recordings associated with different emotional states as shown in Table I. These components played a crucial role in training the facial emotion recognition and audio-based emotion recognition models, allowing Empath.ai to understand emotional cues and provide empathetic responses.

TABLE I: Datasets Used in Training

Module	Dataset Used
Facial Emotion Recognition	AffectNet
Speech Emotion Recognition	TESS
Text-based Emotion Recognition	Crowdflower, GoEmotions, ISEAR, MELD, SemEval-2018

B. Algorithm Development

The algorithm development phase involved developing and optimising Empath.ai's multimodal components, including natural language processing (NLP), facial emotion recognition, and audio-based emotion recognition. This comprehensive approach aimed to enhance the chatbot's understanding of user emotions and generate more empathetic and context-aware responses.

The development process involved iterative refinement, experimentation, and hyperparameter tuning to improve the performance of all components. By integrating the NLP, facial emotion recognition, and audio-based emotion recognition, Empath.ai achieved a comprehensive and context-aware understanding of users' emotions, enabling the chatbot to generate more accurate and empathetic responses.

C. Integration of Models in a MERN App

To bring together the various components of Empath.ai into a cohesive and user-friendly experience, a MERN (MongoDB, Express.js, React.js, Node.js) application was developed. This integration allowed for seamless communication between the different models and provided a platform for users to interact with the chatbot effectively.

The MERN app architecture facilitated the storage and retrieval of user conversations, ensuring a personalised and continuous experience. The MongoDB database served as a reliable storage solution, securely storing user inputs and chatbot responses for future reference and analysis.

The Express.js framework handled the server-side logic, enabling communication with the NLP, facial emotion recognition, and audio-based emotion recognition models. Through well-defined APIs, user inputs were passed to the appropriate models for analysis and processing. The responses generated by the models were then relayed back to the user interface for display.

On the client side, React.js powered the user interface, providing an interactive and intuitive platform for users to communicate with Empath.ai. The interface allowed users to

input their queries, view the chatbot's responses, and visualise emotional cues extracted from facial expressions or speech.

The Node.js runtime environment served as the backbone of the application, facilitating real-time communication between the server and client components. It ensured smooth and efficient data flow, enabling instant feedback and responses from the chatbot.

The integration of the NLP, facial emotion recognition, and audio-based emotion recognition models within the MERN app created a seamless and comprehensive user experience. Users could engage in meaningful conversations with Empath.ai, receive empathetic responses, and benefit from the contextual understanding of their emotions.

D. System Architecture

The Empath.ai system is designed to be a context-aware chatbot that offers emotional support and therapy by leveraging multimodal communication. The system architecture consists of three main components as shown in Fig 1:

- 1) Facial Emotion Recognition Module
- 2) Audio-based Speech Emotion Recognition Module
- 3) Text-based Emotion Extraction Module

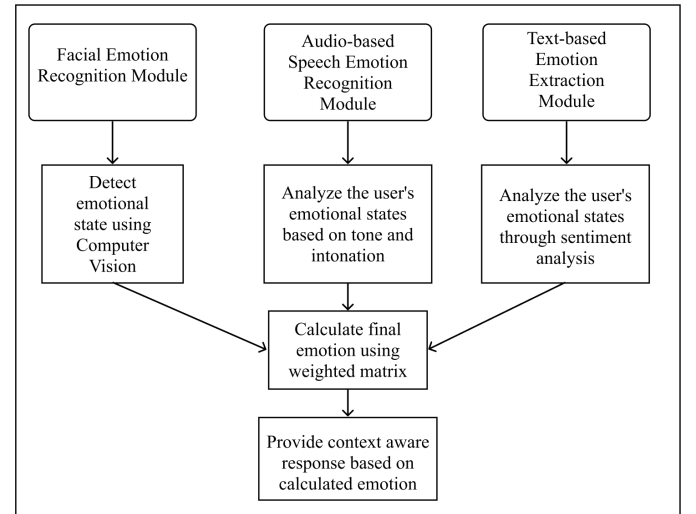


Fig. 1: Empath.ai System Architecture

1) Facial Emotion Recognition Module: We employ multiple neural networks to achieve facial emotion recognition. Firstly, we have a model dedicated to detecting faces in images. This network is a Single Shot MultiBox Detector (SSD) with a MobileNetV1 backbone. It identifies the presence and location of faces in an image. Secondly, we use another network to map facial landmarks, which helps align the detected faces accurately. This step ensures that the subsequent emotion computation is performed correctly. Lastly, we utilise a separate network to compute the emotion based on the aligned facial features, this process is illustrated in Fig. 2. This network is trained on a dataset containing labelled facial expressions, such as neutral, happy, sad, angry, fearful, surprise, and disgust. By using these networks sequentially,

we can extract relevant emotional information and provide context-aware emotional support and therapy as presented in Fig. 3. The training data primarily comes from the AffectNet dataset. It consists of 450,000 facial images labelled with seven emotion categories, including faces captured under different lighting conditions and poses (Mollahosseini et al., 2017) [10].

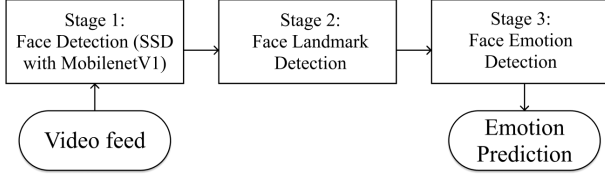


Fig. 2: Pipeline for Facial Emotion Detection



Fig. 3: Output of Stages of Face Module

2) *Audio Based Speech Emotion Recognition Module*: The audio-based speech emotion recognition module is responsible for identifying the emotional state of the user based on their voice by performing audio signal analysis.

In our system, we use a Deep Neural Network (DNN) approach for speech emotion recognition. The DNN is trained on a large dataset of labelled speech samples to classify the emotional state of the user's speech.

For training the machine learning model, we use the Toronto Emotional Speech Set (TESS) dataset [11]. The TESS dataset consists of audio recordings of actors expressing seven basic emotions: anger, disgust, fear, happiness, pleasant surprise, sadness, and neutral. The dataset includes a range of emotions expressed in various vocal intensities, pitches, and styles.

TABLE II: Model Architecture of Audio Module

Layer (type)	Output Shape	Param #
conv1d_2 (Conv1D)	(None, 168, 256)	111616
activation_3 (Activation)	(None, 168, 256)	0
max_pooling1d_2 (MaxPooling 1D)	(None, 21, 256)	0
dropout_2 (Dropout)	(None, 21, 256)	0
conv1d_3 (Conv1D)	(None, 21, 128)	163968
activation_4 (Activation)	(None, 21, 128)	0
max_pooling1d_3 (MaxPooling 1D)	(None, 5, 128)	0
dropout_3 (Dropout)	(None, 5, 128)	0
flatten_1 (Flatten)	(None, 640)	0
dense_2 (Dense)	(None, 64)	41024
dense_3 (Dense)	(None, 7)	455
activation_5 (Activation)	(None, 7)	0

Regarding the model architecture used, we employed a combination of Mel-frequency cepstral coefficients (MFCCs)

and Mel-spectrograms. These techniques aid in extracting essential features from the audio data. These features are then used to classify the user's emotional state into one of several predefined categories. The specific model architecture we implemented is illustrated in Table II.

3) *Text Based Emotion Extraction Module*.: The Text-Based Emotion Extraction Module is a crucial component of the Empath.ai system, designed to extract emotions from text-based input and provide a context-aware chatbot experience. This module is built upon the Emotion English DistilRoBERTa-base model, developed by Jochen Hartmann. [12] The model has been fine-tuned and trained on a diverse range of datasets, enabling it to predict emotions based on Ekman's six basic emotions, including anger, disgust, fear, joy, sadness, and surprise, along with a neutral class.

The training of the DistilRoBERTa-base model involved the utilisation of various datasets. These datasets encompass a wide array of text types, including Twitter data, Reddit posts, student self-reports, and TV dialogue utterances.

To utilise the Text-Based Emotion Extraction Module, the input text is processed through a pipeline involving the following steps:

- 1) *Preprocessing*: The input text is preprocessed to remove punctuation and other unnecessary elements. This step ensures that the model receives clean and standardised text input.
- 2) *Tokenization*: The preprocessed text is tokenized into individual words or sub-word units. Tokenization enables the model to understand the context and meaning of the input text at a granular level.
- 3) *Emotion Classification*: The model classifies the input text into one of the emotion categories: anger, disgust, fear, joy, sadness, surprise, or neutral. The classification is based on the emotional cues and patterns present in the text, learned during the model training phase.

The predicted emotion is extracted from the model's classification output. This extracted emotion can be further processed in tandem with the predictions from the face and audio modules to recognise the user's emotional state. Then with an accurate understanding of the user's emotion the system can generate appropriate responses, tailor the chatbot's behaviour, or provide a more empathetic conversational experience.

E. Weighted Matrix for Final Emotion

After the emotional information is extracted, we suggest using a static weighted matrix with weights based on the reliability of each emotion recognition module and the importance of that mode of communication in the process of conveying emotions. We recommend assigning a weight of 0.5 to the facial emotion recognition module, 0.2 to the audio-based speech emotion recognition module, and 0.3 to the text-based emotion extraction module. The matrix is illustrated in Fig. 4. However, these weights can be adjusted based on the specific use case of the Empath.ai system.

By using a weighted matrix, we can combine the emotional information from multiple sources and generate a more accu-

rate and reliable emotion for the user. In this way the final emotion that is more precise but also truly captures the user's emotional state.

$$\begin{matrix} & \text{Facial} & \text{Audio} & \text{Text} \\ \begin{bmatrix} 0.5 & 0.2 & 0.3 \end{bmatrix} \end{matrix}$$

Fig. 4: Weighted Matrix

F. Deciding Final Emotion

We take the three vectors from the face, audio and text modules and combine them into a matrix. We then multiply this matrix with the weighted matrix to get the final emotion scores. This multiplication process generates new values for each emotion, taking into consideration the assigned weights. By doing so, a more accurate determination of the user's final emotion is achieved, accounting for the contextual factors at play.

The formula used to calculate the final emotion value, denoted as E , is as follows:

$$FinalVector = CombinedMatrix * Weight$$

After applying the weighted matrix to each emotion and obtaining the corresponding values for all seven emotions, the emotion with the highest score is identified. This particular emotion is considered the user's dominant emotion and is returned as the final result.

G. Response Generation

Empath.ai utilises the Facebook Blenderbot-400M-Distill model, integrated into a conversational pipeline, to generate personalised and empathetic responses. By considering the computed emotion of the user, the model leverages deep learning techniques, attention mechanisms, and language modelling to produce natural, contextually relevant, and supportive interactions. This approach ensures that users receive engaging and empathetic responses tailored to their individual needs.

The proposed Empath.ai is a comprehensive system for context-aware emotional support and therapy. It employs facial emotion recognition, audio-based speech emotion recognition, text-based emotion extraction, and text generation. These modules are trained on labelled emotional data using advanced machine learning techniques. The system then computes the contextually nuanced final emotion which helps it when formulating its response to the user.

IV. RESULTS

The evaluation's findings are quite encouraging. Users claimed that the chatbot could recognise their emotional state with accuracy and could offer pertinent and beneficial support.

The accuracy of the facial emotion recognition module surpassed 92% when evaluated on the AffectNet dataset, which serves as the training data for the machine learning model. The audio-based speech emotion recognition module achieved an accuracy of over 90% on the TESS dataset, which was used

to train the machine learning model. The text-based emotion recognition module achieved an accuracy of over 86% on a mixed dataset, which was used to train the machine learning model, the accuracies of the different modules are illustrated in Table III.

TABLE III: Accuracy of Emotion Recognition Modules

Emotion Recognition Module	Accuracy
Facial Emotion Recognition	92%
Speech Emotion Recognition	90%
Text-based Emotion Recognition	86%

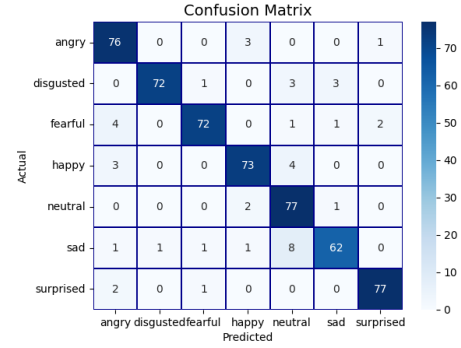


Fig. 5: Face: Confusion Matrix

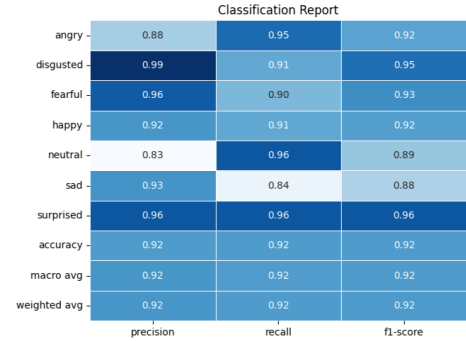


Fig. 6: Face: Classification Report

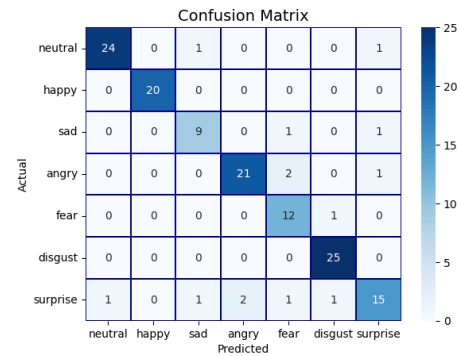


Fig. 7: Audio: Confusion Matrix

neutral	0.96	0.92	0.94
happy	1.00	1.00	1.00
sad	0.82	0.82	0.82
angry	0.91	0.88	0.89
fear	0.75	0.92	0.83
disgust	0.93	1.00	0.96
surprise	0.83	0.71	0.77
accuracy	0.90	0.90	0.90
macro avg	0.89	0.89	0.89
weighted avg	0.90	0.90	0.90
	precision	recall	f1-score

Fig. 8: Audio: Classification Report

	anger	disgust	fear	joy	neutral	sadness	surprise
anger	65	0	0	0	2	0	8
disgust	5	62	1	0	0	5	2
fear	1	2	64	0	3	3	2
joy	6	0	0	65	4	0	0
neutral	0	0	2	3	70	0	0
sadness	4	1	3	0	0	67	0
surprise	5	0	1	0	4	6	59
	anger	disgust	fear	joy	neutral	sadness	surprise

Fig. 9: Text: Confusion Matrix

anger	0.76	0.87	0.81
disgust	0.95	0.83	0.89
fear	0.90	0.85	0.88
joy	0.96	0.87	0.91
neutral	0.84	0.93	0.89
sadness	0.83	0.89	0.86
surprise	0.83	0.79	0.81
accuracy	0.86	0.86	0.86
macro avg	0.87	0.86	0.86
weighted avg	0.87	0.86	0.86
	precision	recall	f1-score

Fig. 10: Text: Classification Report

V. DISCUSSION

The development of Empath.ai as a context-aware chatbot for emotional support and therapy has significant potential to improve the mental health outcomes of individuals worldwide.

The COVID-19 pandemic has served to reinforce the critical role that IVFs, such as Empath.ai, play in promoting and maintaining mental well-being. Studies have shown a significant increase in symptoms of depression and anxiety among individuals worldwide during the pandemic [13], [14]. However, many individuals do not have access to traditional mental health resources due to factors such as cost, stigma,

and limited availability [15]. Chatbots like Empath.ai offer a potential solution by providing a low-cost and accessible option for emotional support and guidance.

The development of Empath.ai also contributes to the growing field of affective computing, which seeks to develop systems and technologies that can recognise, interpret, and respond to human emotions [16]. Empath.ai's use of multimodal emotion recognition techniques and context-aware responses represents a significant advancement in this field.

Overall, the development of Empath.ai represents a significant contribution to both the field of affective computing and the broader goal of improving mental health outcomes worldwide.

VI. CONCLUSION AND FUTURE SCOPE

Empath.ai is an innovative and promising solution that leverages multimodal communication and cognitive-behavioural therapy techniques to provide tailored emotional support and therapy for individuals struggling with mental health issues. Through its advanced emotion extraction modules, Empath.ai can recognise emotional cues from users and respond with appropriate support. It holds the potential to bridge the mental health treatment gap and provide a much-needed solution for those who may not have access to traditional forms of therapy.

Empath.ai offers a personalised experience of emotional support and guidance to individuals grappling with mental and emotional health challenges, encompassing conditions such as anxiety, depression, and stress. Empath.ai creates a non-judgemental and easily accessible space that aids individuals in managing their emotional well-being. With a commitment to convenience, confidentiality, and efficacy, the chatbot strives to enhance the overall quality of life for its users by providing the invaluable resource of emotional support.

Empath.ai revolutionizes emotional support and therapy, offering a personalized and context-aware system.

To enhance Empath.ai's capabilities, we can focus on future development and improvement. One area is expanding its language capabilities to offer multilingual support, reaching a wider audience beyond English speakers. Additionally, incorporating additional modalities like gesture recognition and biometric data analysis would further improve the system's functionality. These advancements would contribute to a more inclusive and comprehensive experience for users.

By integrating human therapists into the system, we enhance its effectiveness. The chatbot system adeptly handles minor cases, freeing up human therapists to prioritize critical and complex cases. This optimized resource allocation ensures efficient support. Moreover, when the chatbot faces confusion or complexities, users benefit from the expertise and guidance of human professionals. This collaboration combines the strengths of both technology and human expertise, resulting in a powerful and comprehensive support system.

After more testing and development, it would be very advantageous to release Empath.ai to a larger audience so that anyone in need of mental health treatment anywhere in the world could use it.

REFERENCES

- [1] World Health Organization. "Mental disorders" WHO. <https://www.who.int/news-room/fact-sheets/detail/mental-disorders> (accessed February. 25, 2023).
- [2] B. Pfefferbaum and C. S. North, "Mental Health and the Covid-19 Pandemic," *New England Journal of Medicine*, vol. 383, no. 6, pp. 510–512, 2020, doi: 10.1056/NEJMp2008017.
- [3] G. Thornicroft, "Most people with mental illness are not treated.," *The Lancet*, 2007.
- [4] A. Peipert, A. C. Krendl, and L. Lorenzo-Luaces, "Waiting Lists for Psychotherapy and Provider Attitudes Toward Low-Intensity Treatments as Potential Interventions: Survey Study," *JMIR Form Res*, vol. 6, no. 9, p. e39787, Sep. 2022, doi: 10.2196/39787.
- [5] K. K. 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 Ment Health*, vol. 4, no. 2, p. e19, Jun. 2017, doi: 10.2196/mental.7785.
- [6] T. W. Bickmore, L. M. Pfeifer, and B. W. Jack, "Taking the Time to Care: Empowering Low Health Literacy Hospital Patients with Virtual Nurse Agents," 2009. doi: 10.1145/1518701.1518891.
- [7] L. Laranjo et al., "Conversational agents in Healthcare: A systematic review," *Journal of the American Medical Informatics Association*, vol. 25, no. 9, pp. 1248–1258, 2018. doi:10.1093/jamia/ocy072
- [8] G. Park, J. Chung, and S. Lee, "Effect of AI chatbot emotional disclosure on user satisfaction and reuse intention for Mental Health Counseling: A serial mediation model - current psychology," *Springer-Link*, <https://link.springer.com/article/10.1007/s12144-022-03932-z> (accessed May 21, 2023).
- [9] S. Hoermann, K. L. McCabe, D. N. Milne, and R. A. Calvo, "Application of synchronous text-based dialogue systems in Mental Health Interventions: Systematic Review," *Journal of Medical Internet Research*, vol. 19, no. 8, 2017. doi:10.2196/jmir.7023
- [10] A. Mollahosseini, B. Hasani and M. H. Mahoor, "AffectNet: A Database for Facial Expression, Valence, and Arousal Computing in the Wild," in *IEEE Transactions on Affective Computing*, vol. 10, no. 1, pp. 18-31, 1 Jan.-March 2019, doi: 10.1109/TAFAC.2017.2740923.
- [11] K. Dupuis and M. Kathleen Pichora-Fuller, "Recognition of emotional speech for younger and older talkers: Behavioural findings from the toronto emotional speech set", *Canadian Acoustics*, vol. 39, no. 3, pp. 182–183, Sep. 2011.
- [12] J. Hartmann. "Emotion English DistilRoBERTa-base", 2022. [Online]. Available: <https://huggingface.co/j-hartmann/emotion-english-distilroberta-base> (accessed Nov 04, 2022).
- [13] M. É. Czeisler et al., "Mental health, substance use, and suicidal ideation during the COVID-19 pandemic—United States, June 24–30, 2020," *Morbidity and Mortality Weekly Report*, vol. 69, no. 32, p. 1049, 2020.
- [14] Salari, N., Hosseini-Far, A., Jalali, R. et al. Prevalence of stress, anxiety, depression among the general population during the COVID-19 pandemic: a systematic review and meta-analysis. *Global Health* 16, 57 (2020). <https://doi.org/10.1186/s12992-020-00589-w>
- [15] J. E. Hollander and B. G. Carr, "Virtually Perfect? Telemedicine for Covid-19," *New England Journal of Medicine*, vol. 382, no. 18, pp. 1679–1681, 2020, doi: 10.1056/NEJMp2003539.
- [16] R. W. Picard, "Affective computing: challenges," *International Journal of Human-Computer Studies*, vol. 59, no. 1, pp. 55–64, 2003, doi: [https://doi.org/10.1016/S1071-5819\(03\)00052-1](https://doi.org/10.1016/S1071-5819(03)00052-1).