AI Therapist

Authors: Saraswati Patil, Aditya Patil, Anuja Deshpande, Surabhi Deshpande, Dev Bohra

***Abstract:** AI Therapist model that utilizes LSTM networks and NLP techniques to deliver personalized mental health support. The model incorporates emotion detection, customized solutions, song recommendations, and emotion tracking capabilities. By analyzing natural language input, the AI Therapist accurately identifies users' expressed emotions and provides empathetic responses. Integration with the OpenAI API enhances the model's language capabilities, allowing for detailed explanations and practical guidance. The AI Therapist also tracks and logs users' emotions over time, enabling personalized interventions based on emotional patterns. Evaluation results demonstrate the effectiveness of the AI Therapist in providing meaningful mental health support. This comprehensive approach to mental healthcare empowers individuals to improve their emotional well-being.*

Keywords—Artificial Intelligence, emotions, LSTM, mental health, NLP, Open AI, therapist.

I. Introduction

In the realm of mental healthcare, there is an increasing need for innovative and accessible therapy solutions. Leveraging the power of Natural Language Processing (NLP) and Artificial Intelligence (AI), we have embarked on a research project to develop an AI Therapist model. This model harnesses the capabilities of Long Short-Term Memory (LSTM) and NLP techniques to offer a comprehensive set of features aimed at supporting individuals' mental well-being.

Our AI Therapist model goes beyond simple chatbot interactions by incorporating advanced functionalities. It includes emotion detection, enabling the model to discern and understand the underlying emotions expressed in user input. Drawing upon the OpenAI API, our model provides personalized and tailored responses to users' queries, ensuring a more effective therapeutic experience.

Furthermore, our AI Therapist model offers song recommendations as an additional therapeutic tool. Music has long been recognized for its potential to evoke emotions and provide comfort, making it a valuable resource for mental health support.

To enhance the effectiveness of therapy, our model keeps a log of users' emotions throughout the month. This allows for a comprehensive understanding of emotional patterns and facilitates the provision of targeted interventions and support.

By combining cutting-edge AI techniques, NLP, and the integration of emotional understanding, personalized responses, and music recommendations, our AI Therapist model aims to revolutionize mental healthcare delivery. We anticipate that our research will contribute to the advancement of AI-assisted therapy, providing individuals with accessible, scalable, and tailored support for their mental well-being.

II. Literature review

[1]The paper highlights the increasing occurrence of mass shootings in the United States and their detrimental effects on individuals' mental health. It mentions the statistics of shootings over the years and the negative psychological outcomes experienced by victims, communities, and the public. The authors refer to previous literature indicating the strong relationship between mass shootings and mental health issues such as fear, depression, and PTSD. A survey was conducted by the authors in July 2019 with 1,114 participants who had used chatbot services. The survey measured participants' motivations, protection motivations, active communicative actions, and engagement behaviors related to healthcare chatbot use. Participants showed moderately high motivations and perceived protection, as well as moderately low online and offline engagement behaviors.

[2] The paper examines the application of artificial intelligence (AI) in mental health care, focusing on personal sensing, natural language processing (NLP), and chatbots. Personal sensing technologies like smartphones and wearables can collect data on behaviors and physiological signals, enabling AI to analyze patterns and identify mental health indicators. NLP techniques, such as sentiment analysis and linguistic style analysis, can assess mental well-being through language patterns. Chatbots powered by AI offer scalable and accessible mental health support. They engage in interactive conversations, providing emotional support and therapeutic interventions. Examples of AI-based chatbot platforms include Woebot, an automated conversational agent for mood tracking and cognitive-behavioral therapy techniques, and Tess, a text-based mental health coach. Ethical considerations are emphasized, including privacy, data security, and transparency of AI algorithms. Human involvement and professional oversight are also highlighted to avoid overreliance on technology. Overall, this review paper offers insights into the potential benefits and challenges of AI in mental health care, with personal sensing, NLP, and chatbots as promising avenues for improving mental health outcome

[3] The paper discusses the use of artificial intelligence (AI) in mental healthcare. It explores the clinical applications of AI, barriers, facilitators, and the concept of artificial wisdom. The authors highlight the potential of AI in predictive modeling, early detection, treatment optimization, and remote monitoring. They also address the ethical considerations, privacy concerns, and regulatory frameworks associated with AI implementation. The paper emphasizes the importance of combining AI algorithms with human expertise to achieve "artificial wisdom" in clinical decision-making. It mentions the use of chatbots as examples of AI-based interventions in mental healthcare,

offering scalable and accessible support through automated conversational agents. The overall focus is on the benefits and challenges of integrating AI into mental healthcare and the need for responsible and effective AI tools.

[4] The paper explores the role of machine learning and natural language processing in the field of mental health. The authors provide an overview of the applications of these techniques, ranging from automated diagnosis and risk assessment to predicting treatment outcomes and monitoring mental health through digital data. They discuss the potential benefits, such as improved accuracy, efficiency, and accessibility of mental healthcare, but also highlight several challenges that need to be addressed. These include the need for standardized datasets, robust algorithms, and ethical considerations regarding privacy and data security.

The paper emphasizes the importance of interdisciplinary collaboration between mental health professionals, data scientists, and technologists to develop effective and reliable machine learning models. It also underscores the potential of these technologies to provide personalized and timely interventions, early detection of mental health issues, and support for clinical decision-making. The authors advocate for further research and development in this area to advance the field of mental health and leverage the power of artificial intelligence for the benefit of patients. Overall, this review highlights the promising applications of machine learning and natural language processing in mental healthcare while acknowledging the need for careful implementation and ongoing evaluation.

[5] This paper review examines the role of chatbots and conversational agents in the field of mental health, shedding light on their potential applications in supporting individuals with mental health conditions. The paper discusses how these technologies can assist in screening, assessment, and monitoring of mental health symptoms, offering a promising avenue for early intervention and proactive care. It also explores the challenges associated with chatbots, such as the limitations of natural language processing and the importance of maintaining a human-centered approach in mental healthcare delivery. The review emphasizes the need for rigorous evaluation and validation of chatbots' effectiveness and safety, as well as the importance of addressing ethical considerations, such as maintaining user confidentiality and ensuring informed consent. Ultimately, the paper highlights the evolving landscape of mental health interventions and the potential of chatbots and conversational agents to augment and enhance traditional approaches to care.

[6] This research paper focuses on emotion detection from text documents and explores the techniques used in this domain. Emotion detection is a content-based classification problem that involves Natural Language Processing and Machine Learning concepts. The paper emphasizes that emotions can be expressed through various means, including facial expressions, gestures, speech, and written text. The authors discuss the recognition of emotions based on textual data and delve into the methods employed in emotion detection. The study highlights the importance of combining techniques from Natural Language Processing and Machine Learning to effectively analyze and classify emotions in text. The aim is to develop models and algorithms that can accurately identify and categorize emotions expressed in written text. By addressing this problem, the research contributes to the understanding and advancement of emotion detection technology, which can have applications in fields such as sentiment analysis, mental health assessment, and human-computer interaction.

III. TOOLS AND TECHNOLOGIES

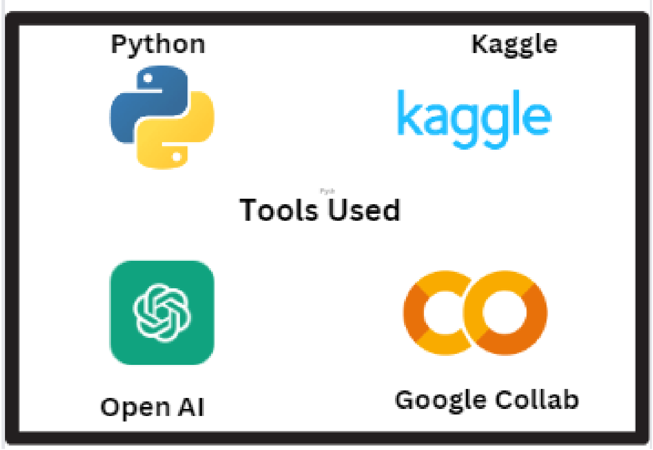


Fig.1 Tools Used

A. GloVe: Global Vectors for Word Representation

GloVe (Global Vectors for Word Representation) is a word representation model introduced in [7]. It is a method for learning vector representations of words that capture semantic relationships between them.

The GloVe model combines the advantages of both global and local word representation approaches. It leverages the co-occurrence statistics of words within a large corpus to learn vector representations that encode semantic meaning. Unlike traditional local context-based models like Word2Vec, which focus on predicting the next word in a sequence, GloVe considers the overall co-occurrence patterns of words in a corpus.

The key idea behind GloVe is that word vectors can be represented as ratios of co-occurrence probabilities. By analyzing the statistics of word co-occurrences, GloVe learns vector representations that capture meaningful relationships between words. This allows the model to capture not only syntactic similarities but also semantic relationships between words.

GloVe has been widely adopted in natural language processing tasks and applications. Its pre-trained word vectors have proven to be valuable in various downstream tasks such as sentiment analysis, machine translation, and question-answering systems. The model's ability to generate high-quality word representations has contributed to significant advancements in natural language understanding and has become an essential tool for researchers and practitioners in the field.

B. LTSM

The Bidirectional LSTM (Long Short-Term Memory) is a type of recurrent neural network (RNN) architecture that consists of two LSTM layers: one that processes the input sequence in the forward direction, and another that processes it in the backward direction. This architecture allows the model to capture both past and future context when making predictions.

In a standard LSTM, information flows only from past to future, meaning that the predictions at each time step depend only on the previous time steps. However, in tasks where future context is also important, such as natural language processing, the Bidirectional LSTM proves to be beneficial.

By using a Bidirectional LSTM, the model can take advantage of the forward LSTM to capture past dependencies and the backward LSTM to capture future dependencies. This allows the model to have a more comprehensive understanding of the input sequence and make more accurate predictions.

The output of the Bidirectional LSTM is typically obtained by combining the outputs from both the forward and backward LSTMs, either through concatenation or an element-wise operation.

Bidirectional LSTMs have been widely used in various tasks such as speech recognition, sentiment analysis, machine translation, and named entity recognition. They are particularly effective when the context in both directions is crucial for the task at hand.

C. Dataset

Emotion detection from text presents a significant challenge in Natural Language Processing due to the lack of labeled datasets and the multi-class nature of the problem. The availability of a labeled dataset specifically for emotion detection allows us to tackle this challenge. The objective is to develop an effective model that can accurately identify emotions in text. The dataset comprises tweets that have been annotated with corresponding emotions. The data includes three columns: tweet_id, sentiment, and content. The "content" column contains the raw tweets, while the "sentiment" column indicates the emotion expressed in the tweet. Addressing the issue of class imbalance caused by the variety of human emotions is essential. For more details and a deeper understanding, refer to the provided starter notebook.

D. OpenAI API

The OpenAI API is a tool that allows developers to access and utilize OpenAI's language models, including ChatGPT, in their own applications. It provides an easy way to integrate powerful natural language processing capabilities into your software. With the API, you can make HTTP requests to the OpenAI endpoint and receive responses in JSON format. It offers fine-grained control over the model's behavior, allowing you to customize parameters and provide instructions to guide the output. The API has a usage-based pricing model, where you pay based on the number of tokens processed. OpenAI provides a Python library to simplify the integration process. It's important to refer to the OpenAI documentation for specific guidelines, model availability, usage limits, and pricing details. The OpenAI Playground is a useful resource for testing and experimenting with the API.

IV. Methodology

- Perform data preprocessing on the dataset – remove null values, duplicate values; data visualization.
- Split the dataset into train, test and validation dataset.
- Perform data cleaning – remove urls, punctuation, convert to lowercase, remove extra whitespaces etc.
- Compare performance of various models – SVM, RF, Logistic Regression, Decision Trees.
- Perform label encoding, tokenize words, pad to make equal length words.
- Build Bidirectional LSTM Model since it has been found to best preserve context from past and future.
- Train the model until loss keeps decreasing.
- Test input data and get emotion from the text.
- Store all emotions along with timestamp in a csv file.
- By integrating OpenAI’s API, send instant suggestions on how to handle negative emotions.
- Build a song-recommendation system by creating an appropriate dataset for all emotions present – sadness, anger, fear, joy, surprise, love.

V. SYSTEM ARCHITECTURE

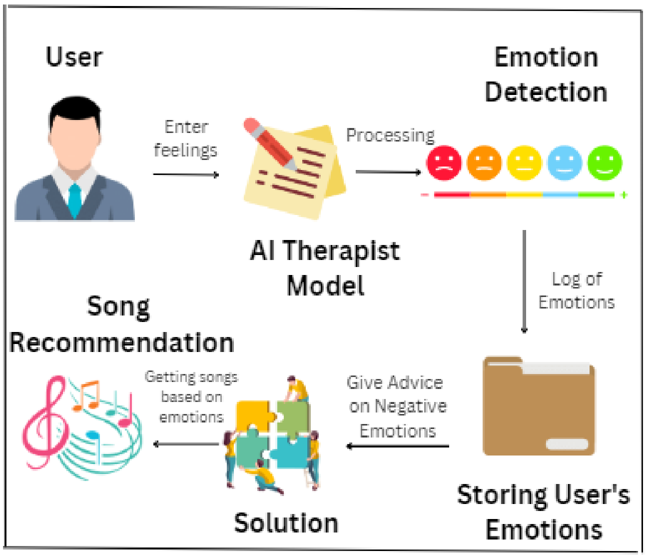


Fig.2 Workflow Diagram

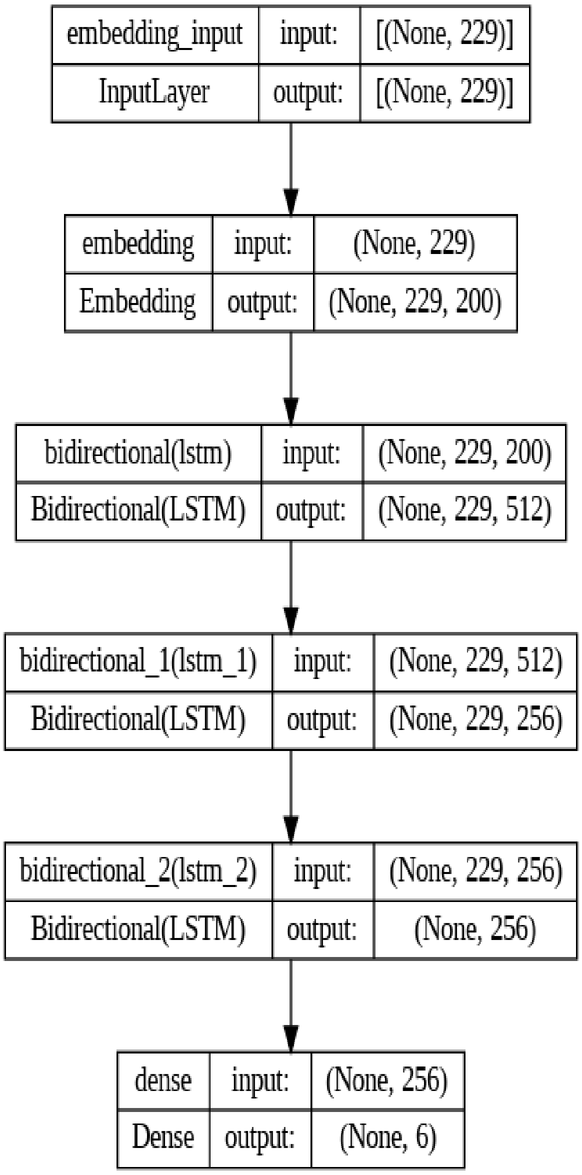
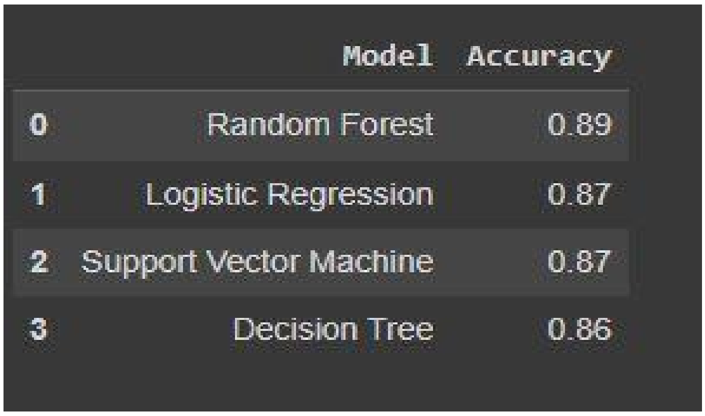


Fig.3 Model Architecture Diagram

VI. RESULTS AND DISCUSSIONS

We have successfully implemented the AI Therapist Model. Here are few pictures of our project.



	Model	Accuracy
0	Random Forest	0.89
1	Logistic Regression	0.87
2	Support Vector Machine	0.87
3	Decision Tree	0.86

Fig.4 Comparison of Various Models

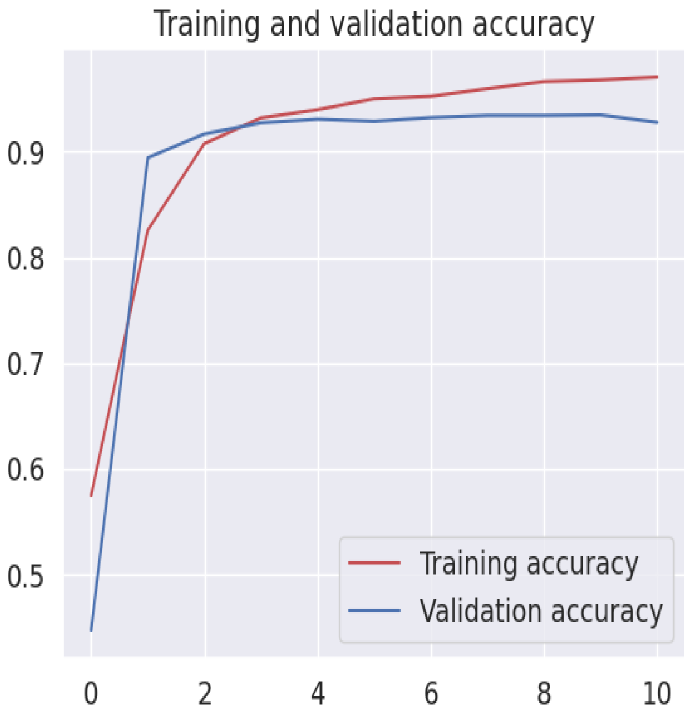


Fig.5 Training and Validation accuracy for Bidirectional LSTM Model

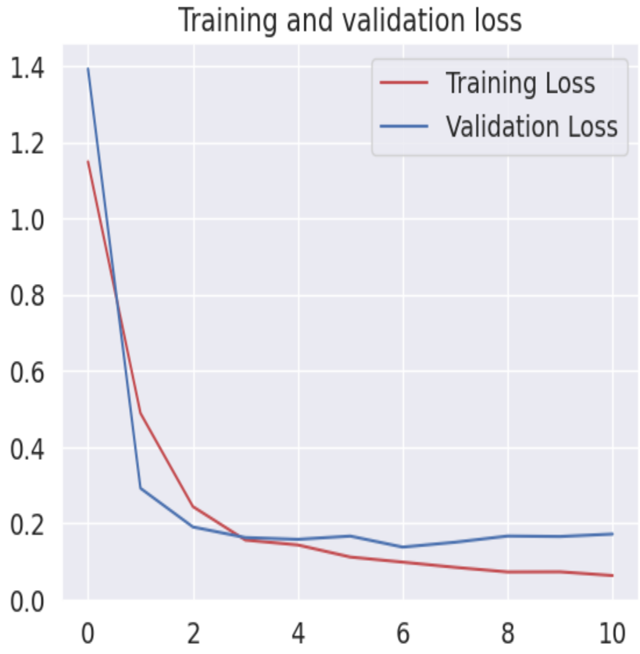


Fig.6 Training and Validation loss for Bidirectional LSTM Model

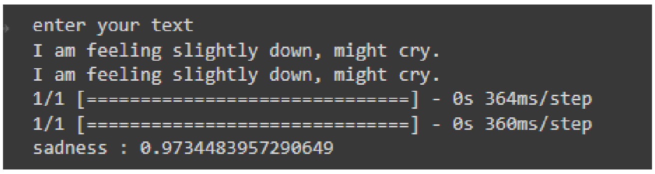


Fig.6 Emotion Classification

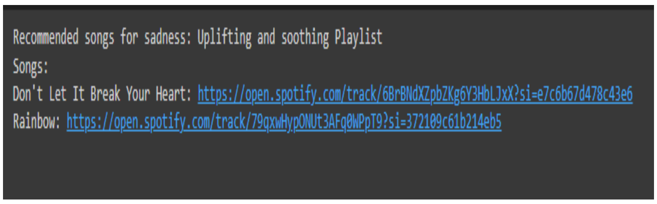


Fig.7 Song Recommendation

VII. CONCLUSION

The AI Therapist model is developed using the bidirectional LSTM. This model gives an accuracy of 93%.AI Therapist will detect the user's emotions based on the responses. It will suggest appropriate solutions to the problems.AI Therapist will also record logs of the user's emotions throughout the month, this data will be useful for further analysis.AI therapist can also recommend songs based on the user's mood. This model will be helpful to psychiatrists and therapists for curing their clients. Thus, AI Therapist will help user's to combat depression, anxiety, etc.

VIII. FUTURE SCOPE

The future of AI therapists using NLP also involves incorporating advanced sentiment analysis and emotion recognition techniques to better understand and respond to users' emotional states. This could enable AI therapists to provide more empathetic and tailored support. Furthermore, advancements in AI and NLP technologies may lead to the development of virtual reality (VR) or augmented reality (AR) applications, allowing users to engage in immersive therapy experiences. Finally, the integration of AI therapists with smart home devices and virtual assistants may enable continuous support and seamless integration into individuals' daily lives.

References

[1] Yang Cheng,Hua Jiang, AI-Powered mental health chatbots: Examining users' motivations, active communicative action and engagement after mass-shooting disasters,Wiley,2020.

[2] Simon D'Alfonso,AI in mental health,Current Opinion in Psychology,Volume 36,2020,Pages 112-117,ISSN 2352-250 Lee E.E., Torous J., De Choudhury M., Depp C.A., Graham S.A., Kim

[3] H.-C., Paulus M.P., Krystal J.H. & Jeste D.V., Artificial Intelligence for Mental Healthcare: Clinical Applications, Barriers, Facilitators, and Artificial Wisdom, Biological Psychiatry: Cognitive Neuroscience and Neuroimaging (2021), doi: https://doi.org/10.1016/j.bpsc.2021.02.001.X,https://doi.org/10.1016/j.copsyc.2020.04.005.

[4] Aziliz Le Glaz, Yannis Haralambous, Deok-Hee Kim-Dufor, Philippe Lenca, Romain Billot, et al..Machine Learning and Natural Language Processing in Mental Health: Systematic Review. Journal of Medical Internet Research, JMIR Publications, 2021, 23 (5), pp.e15708. ff10.2196/15708ff. ffhal03217684f

[5] Aditya Nrusimha Vaidyam, BS1 , Hannah Wisniewski, BS1,John David Halamka, MD1, Matcheri S. Kashavan1, and John Blake Torous, MD, MB, Chatbots and Conversational Agents in Mental Health: A Review of the Psychiatric Landscape, The Canadian Journal of Psychiatry,2020

[6] Shivhare, Shiv Naresh & Khethawat, Saritha. (2012). Emotion Detection from Text. Computer Science & Information Technology. 2. 10.5121/csit.2012.2237.

[7] Jeffrey Pennington, Richard Socher, and Christopher Manning. 2014. GloVe: Global Vectors for Word Representation. In Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP), pages 1532–1543, Doha, Qatar. Association for Computational Linguistics.

[8] Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. Neural computation, 9(8), 1735-1780.