International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 08 Issue: 06 | June - 2024 SJIF Rating: 8.448 ISSN: 2582-3930

AI Driven Chatbot Counsellor

Prof. Rekha Kulkarni
Department of Computer Engineering
SCTR'S Institute of Computer Technology , Dhankawadi, Pune
rakulkarni@pict.edu

Ruchita Amale
Department of Computer Engineering
SCTR'S Institute of Computer Technology
Sangamner, Ahmednagar
amaleruchita@gmail.com

Mrudula Jambhalkar
Department of Computer Engineering
SCTR'S Institute of Computer Technology
Kalyan, Thane
mrudulavj2001@gmail.com

Aaryan Dharamadhikari
Department of Computer Engineering
SCTR'S Institute of Computer Technology
Dhankawadi, Pune
dharmadhikariaaryan@gmail.com

Abstract - Speech and written information are fundamental to human communication. As a result, the majority of spoken and textual communication takes place on digital platforms like Twitter, Facebook, and WhatsApp, among others. Our model employs dual recurrent neural networks (RNNs) to encode the information from text and audio sequences, as spoken language and sound constitute emotional discourse. The emotion class is then predicted by combining the data from the two sources. Due to the complexity of speech emotion recognition, models that use audio properties to generate powerful classifiers have become increasingly important. Completing sentences or crafting sentences from a given starting word is a major aspect of natural language processing. In a way, it shows if a machine is capable of human creativity and mental processes. In order to assist handle diverse phrase generation challenges, we use natural language processing to train the machine for certain tasks. This is especially useful for application situations such as machine translation, automatic question answering, and summary creation. At the moment, OpenAI GPT and BERT are the most widely used language models for text generation and prediction. The approaches based on handwritten instructions, patterns, or statistical methods have been quickly superseded by the latest developments in deep learning and artificial intelligence, such as end-to-end trainable neural networks. This research presents a novel approach to deep neural learning-based chatbot creation. This approach builds a multilayer neural network to analyse and learn from the data. Furthermore, we employ supplementary limitations on the generation model to generate the correct response, which is capable of discerning the context of the discussion, the user's mood, and the anticipated response. This enables us to provide customised counselling replies depending on customer feedback. Through this study, two new corpora will be used to train the OpenAI GPT model, which will then be used to generate articles and long sentences. Finally, a comparison study will be carried out. Concurrently, we will use the BERT model to complete the task of context-based intermediate word prediction.

Keywords : Artificial Intelligence, Data Science, NLP, Deep Learning, Machine Learning, GPT, Generative AI, Speech Synthesis

I. INTRODUCTION

The identification of human emotions has long been a subject of study. Artificial intelligence (AI) techniques are an appropriate strategy, according to several recent research. The research employs a variety of emotionally labelled data to construct different emotion classification models. Chatbots are computer programs that are typically made to realistically mimic human behaviour in order to help customers by acting as a chat partner. The job of natural language processing known as natural language generation (NLG) involves creating natural language using a logical form from a machine representation system, such a knowledge base. One may compare an NLG system to

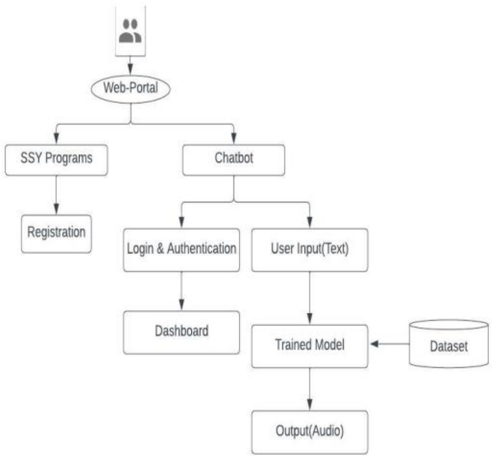
a translator that transforms data into a representation in natural language. And hence when this type of chatbot would be able to recognize the human emotions then counselling them will help patients in terms of their medical issues. One disadvantage of traditional chatbot systems is that their question-answer pairings are predefined in the database, so they always provide the same response for every client inquiry. When a consumer notices that a machine is responding to their query, they lose interest in the chatbot since it always provides the same answer. By generating variants of a predefined response using natural language generation, humanising the chatbot will give it a more human touch and personal touch. We'll employ strategies like paraphrasing to create variations and boost client satisfaction by having the response sound more genuine. Predicting the emotional content of speech and categorising speech based on one or more labels—such as happy, sad, neutral, or angry—are the objectives of speech emotion recognition. Although several deep learning techniques have been used to improve emotion classifier performance, this job is still regarded as difficult for a number of reasons. First, the costs of human involvement mean that there is not enough data available to train large neural network-based models. Second, low-level speech cues must be used to teach emotions and their traits. Using feature-based models for this situation shows their limitations.A variety of paraphrase strategies and machine learning algorithm combinations will be employed. Improving the current chatbot in any way to increase customer happiness will be the goal. One such area of natural language creation that focuses mostly on semantics is paraphrase generation. A single statement can be paraphrased by expressing it in several ways while maintaining its meaning.

II. LITERATURE SURVEY

The creation of chatbots for use in education is examined in the study "Chatbot: An automated conversation system for the educational domain" by Anupam Mondal et al. It talks about implementation, design, and advantages including administrative support and individualised learning. It tackles issues and seeks to optimise educational procedures while acting as a resource for educators and developers. Dongkeon Lee, Kyo-Joong Oh, and Ho-Jin Choi's article "The ChatBot Feels You – A Counseling Service Using Emotional Response Generation" probably focuses on developing a counselling service that employs chatbots that can provide emotionally responsive exchanges. In order to improve the user experience in counselling sessions, this research may investigate the incorporation of emotional intelligence into chatbots. The paper "A Deep Neural Network Based Human to Machine Conversation Model" by G Krishna Vamsi and Gaurav Hajela presents research on creating a chatbot utilising deep neural networks for human-machine interactions. It likely explores the architecture, training, and performance of the model, aiming to enhance the naturalness and effectiveness of conversational exchanges between users and machines. The Centre for Spoken Language Understanding (CSLU) at Oregon Graduate Institute's Alexander Kain and Michael W. Macon describe in their paper "Spectral Voice Conversion for Text to Speech Synthesis" a voice conversion technique that uses Gaussian mixture models to change the speech of a source speaker into the voice of a target speaker. Assessments show that it is effective at synthesising text to speech. The manuscript "A Predictive and Generative Text System: Initial Training on Novel Corpora" Using BERT and GPT-2" by Yuanbin Qu et al. looks into the use of language models that have already been trained, including BERT and GPT-2, for tasks involving text production. It assesses their effectiveness and provides insights into their performance, especially in the areas of extended sentence production and masked word prediction. The paper "A Proposed Chatbot Framework for COVID-19" by Eslam Amer, Ahmed Hazem, Omar Farouk, Albert Louca, Youssef Mohamed, and Michel Ashraf introduces a chatbot framework for COVID-19. Utilising BERT-based question-answering, it categorises inquiries and delivers precise responses, aiming to provide efficient information dissemination and support during the pandemic.

International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 08 Issue: 06 | June - 2024 SJIF Rating: 8.448 ISSN: 2582-3930

III. BLOCK DIAGRAM



IV. METHODOLOGY

In NLP, Whisper for voice-to-text conversion, and the GPT-2 and BERT models are the main techniques used in the creation of an AI-driven counselling chatbot. First, a variety of preprocessed datasets including text and audio samples from counselling sessions are gathered. Preprocessing, tokenization, and cleaning are applied to text data, while feature extraction and format conversion are performed on audio data. Next, the chatbot leverages NLP techniques for sentiment analysis, emotion recognition, and context awareness. The GPT-2 model, known for its generative capabilities, is employed for generating natural and contextually relevant responses. It is fine-tuned specifically for counselling-related conversations, incorporating transfer learning if applicable. Additionally, BERT model is utilised for text generation and prediction tasks, pre-training on new corpora to enhance performance. Algorithms for converting speech to text are incorporated to assess the emotional content of user speech. Features such as whisper and bark recognition are included to capture subtleties in voice expression. These algorithms help recognize the feelings of the user and offer relevant replies. The performance of the chatbot is evaluated using criteria including response coherence, emotional intelligence, and user happiness. Comprehensive testing and iterative refinement based on input from domain experts and users are carried out to improve the efficacy of the chatbot in offering users who are in need of emotional support tailored support.

Sr.No	Model	Used for
1.	GPT-2 (Generative Pre-trained Transformer 2)	Used for text generation and prediction tasks.
2.	Spectral Voice Conversion Models	Utilised Gaussian mixture models for voice conversion.
3.	BERT (Bidirectional Encoder Representations from Transformers)	Employed for context-based word prediction and question-answering tasks.
4.	Whisper	Utilised in spectral voice conversion models for modifying source speaker's speech to sound like a target speaker in text-to-speech synthesis.

International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 08 Issue: 06 | June - 2024 SJIF Rating: 8.448 ISSN: 2582-3930

Sr. no	Library	Used for
1.	TensorFlow	Likely used for implementing deep learning models like GPT-2, spectral voice conversion models, and BERT.
2.	PyTorch	Possibly used for training deep learning models, although not explicitly mentioned in the provided data
3.	NLTK (Natural Language Toolkit)	Possibly used for natural language processing tasks like text tokenization and syntactic analysis

V. ALGORITHM

1. Data Collection and Preprocessing

Collect diverse datasets containing text and audio samples from counselling sessions.

Preprocess the text data by cleaning, tokenization, and removing noise. Convert the audio data into suitable formats and extract features such as MFCCs.

2. NLP Techniques Integration

Implement sentiment analysis, emotion recognition, and context awareness using NLP techniques. Fine-tune the GPT-2 model for generating contextually relevant responses in counselling scenarios. Pre-train the BERT model on new corpora to improve text generation and prediction tasks.

3. Voice-to-Text Conversion

Integrate voice-to-text conversion algorithms to analyse the emotional content of user speech. Implement specialised features like whisper detection and bark detection to capture nuances in voice expression.

4. Algorithm Implementation

Develop algorithms to process user inputs, including text and audio, using NLP and voice-to-text conversion techniques. Utilise the GPT-2 and BERT models to generate personalised responses based on the analysed input.

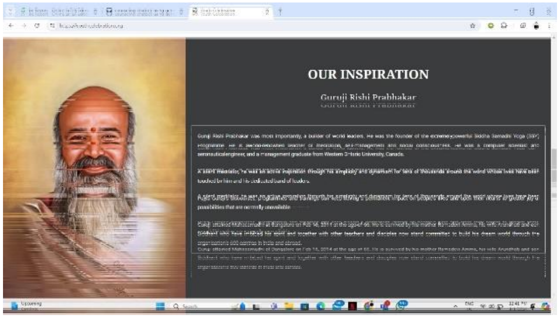
5. Evaluation and Refinement

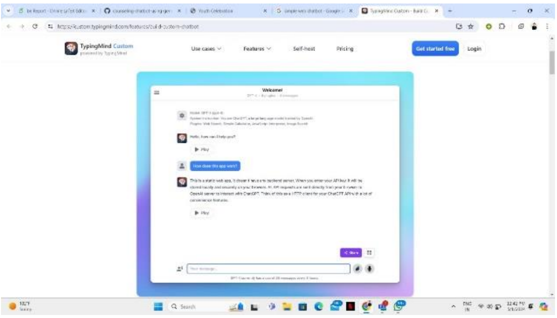
Define evaluation metrics such as response coherence, emotional intelligence, and user satisfaction. Conduct extensive testing to assess the chatbot's performance. Gather user feedback and domain expert input for iterative refinement of the chatbot's algorithms and responses.

6. Deployment and Monitoring:

Deploy the chatbot in real-world counselling scenarios or online platforms. Monitor the chatbot's interactions and performance continuously. Implement mechanisms for updating the chatbot's models and algorithms based on ongoing feedback and improvements.

VI. Output





VII. Conclusion

In summary, the goal of this research is to show the viability and promise of modelling a person's mental processes. To do this, a prepared dataset will be trained using a suitable GPT model using text from books and audio recordings. It highlights how important ethical concerns are at every stage of the project, especially when it comes to protecting privacy and data protection.

VIII. REFERENCES

[1] Anupam Mondal, Monalisa Dey, Dipankar Das , Sachit Nagpal, Kevin Garda **"Chatbot: An automated conversation system for the educational domain" 2018 IEEE**

[2]Dongkeon Lee, Kyo-Joong Oh, Ho-Jin Choi **The ChatBot Feels You – A Counseling Service Using Emotional Response Generation"** 2017 IEEE

[3]G Krishna Vamsi , Gaurav Hajela **A Deep Neural Network Based Human to Machine Conversation Model** G Krishna Vamsi Gaurav Hajela 2020

[4] Mayuresh Virkar ,Vikas Honmane **Humanising the Chatbot with Semantics based Natural Language Generation 2019**

[5] Alexander Kain and Michael W. Macon* Centre for Spoken Language Understanding (CSLU) Oregon Graduate Institute of Science and Technology **Spectral Voice Conversion for Text to Speech Synthesis 1998**

[6]Yuanbin Qu, Peihan Liu, Wei Song, Lizhen Liu*, Miaomiao Cheng Information and Engineering College, Capital Normal University, Beijing **A Text Generation and Prediction System: Pre-training on New Corpora Using BERT and GPT-2 2020**

[7] Seunghyun Yoon, Seokhyun Byun, and Kyomin Jung Dept. of Electrical and Computer Engineering, Seoul National University, Seoul, Korea **Multimodal Speech Emotion Recognition using Audio and Text 2020 IEEE**

[8]Eslam Amer Ahmed Hazem Omar Farouk Albert Louca Youssef Mohamed Michel Ashraf **A Proposed Chatbot Framework for COVID-19 2021 IEEE**

[9]Jia-Chen Gu1, Tianda Li2, Quan Liu1,3, Zhen-Hua Ling1, Zhiming Su3, Si Wei3, Xiaodan Zhu2 1National Engineering Laboratory for Speech and Language Information Processing, University of Science and Technology of China, **Speaker-Aware BERT for Multi-Turn Response Selection in Retrieval-Based Chatbots 2020**

[10]Ruth M. Reeves a,d,*, Lee Christensen g, Jeremiah R. Brown e, Michael Conway g, Maxwell Levis e, Glenn T. Gobbel a,c,d, Rashmee U. Shah f, Christine Goodrich e, Iben Ricket e, Freneka Minter a, Andrew Bohm e, Bruce E. Bray f,g, Michael E. Matheny a,b,c,d, Wendy Chapman **Journal of Biomedical Informatics 2021**