Virtual Psychiatrist

Group Name: Group-6 Transformers

Group Members:

First name	Last Name	Student number
Dev	Makwana	C0885064
Krina	Patel	C0886861
Mahaveersinh	Chauhan	C0884854
Nikita	Maru	C0882340
Trushna	Patel	C0886910

Submission date: 19/04/2024

Table of Contents

List of Figures

Abstract

The provision of mental health support has undergone a revolution in recent years due to the integration of artificial intelligence (AI) with healthcare. Chatbots that utilize natural language processing (NLP) have emerged as potent resources for providing prompt assistance. This research investigates the creation, implementation, and assessment of a GPT-2 focused chatbot designed with psychological concerns in mind. Utilizing extensive textual data from Hugging Face and other sites, the chatbot is trained on a variety of contextually relevant content and can efficiently respond to user requests. This article describes a systematic technique that includes data extraction, cleaning, visualization, training, and assessment methodologies. The use of metrics like DistilBERT Cosine Similarity to evaluate language coherence and semantic understanding is part of this process. Furthermore, user-friendly interfaces, accessible communication, and effective data management are guaranteed by the integration of ReactJS for front-end development, FastAPI for backend services, and MongoDB for data storage. These factors enhance transparency and effectiveness in the mental health care industry. "Harmony" is the name of the chatbot. Its scope includes providing coping mechanisms, stress, anxiety, and depression therapies, as well as acting as a point of contact for emergency support and mental health services. Providing ongoing support and monitoring enhances traditional therapy.

Introduction

In recent years, artificial intelligence (AI) has blended with healthcare to disrupt how mental health support is offered potentially. Among these technologies, chatbots using natural language processing (NLP) are considered a potent and convenient resource that offers immediate assistance and direction to people with mental health problems. This project extensively explores designing, deploying, and assessing a GPT-2 focused chatbot specifically designed for supporting psychological issues. By using readily available vast volumes of textual data, especially from platforms like Hugging Face, training can be done on diverse content that is also contextually relevant, thus ensuring that they can attend to different inquiries from users.

This report briefly presents a systematic procedure for scaffolding and assessing mental health chatbot using Hugging Face’s datasets. First, the report describes the work methodology, which comprises the steps of data extraction, cleaning the raw data, visualization, training model, and evaluation. A training data set includes greetings and some chatbot details that are added to improve the user's conversation experience. It surpasses simple implementation by far, though, and concentrates on rigorous evaluation methodologies to determine the effectiveness and performance of the chatbot developed. To assess the linguistic coherence, relevance, and semantic understanding of the chatbot in question, metrics such as BLEU score, ROGUE score, and DistilBERT Cosine Similarity are used, which provide insight into its functionality and areas where it could be enhanced.

It also tackles the integration of ReactJS for front-end development, FastAPI for backend services, and MongoDB for data storage and management. These technologies are about making user-friendly interfaces in terms of user interface, accessible communication in the backend and ensuring that data is well-stored and managed, creating a holistic system through which users can get mental health support. The chatbot is named “Harmony”, which seeks to boost client commitment and expedite data processing, thus contributing to improved openness and efficiency of AI-powered chatbots combined with state-of-the-art web technologies in the mental healthcare industry.

Scope

Chatbot for mental health has many applications and covers many facets of supporting mental health. This chatbot offers tools, coping strategies, and individualized interventions to help with the management of stress, anxiety, depression, and other mental health issues. It can also act as the initial point of contact for people looking for information about mental health services or emergency support. Furthermore, by evaluating progress, providing constant assistance between sessions, and providing continuing monitoring, it may be able to supplement traditional therapy.

Applications

Mental health-related chatbots have emerged as valuable tools for providing support, guidance, and resources to individuals dealing with mental health issues. Here are some applications of mental health-related chatbots:

1. **24/7 Support:** It offers round-the-clock support to individuals experiencing mental health crises or needing immediate assistance. It can provide a listening ear and offer coping strategies when going through a bad phase in life.
2. **Stress Reduction:** It offers mindfulness exercises, breathing techniques, guided imagery, and other strategies to help users relax and manage stress more effectively.
3. **Lifestyle Modification:** It can guide adopting healthy lifestyle habits, such as regular exercise, nutrition, and sleep hygiene to improve overall mental well-being.
4. **Anonymous Communication:** It provides a safe and confidential space for individuals to express their thoughts, feelings, and concerns without fear of judgment. This anonymity encourages users to seek help and share their experiences more openly.

Dataset Details

Conversational data about mental health, such as conversations, questions, and answers between people seeking mental health help and peers or mental health specialists, make up the Mental Health related datasets. Numerous subjects are covered, including self-care, eating disorders, sleep disorders, anxiety, depression, and stress management. It also includes a few conversational chat samples produced for testing and study. In addition, transcripts of counseling sessions between clients and mental health practitioners—such as psychologists, counselors, and therapists—can be found in certain datasets. It offers insightful information about therapy relationships, methods of treatment, issues raised by clients, and therapeutic results.

The table below depicts the list of all the datasets used as a part of this project:

No	Dataset Name	Description	Conversations
1	NART-100k Synthetic Dataset *(Jerryjalapeno/Nart-100k-synthetic · Datasets at Hugging Face, 2001)*	A synthetic dataset containing a chat between humans and gpt related to mental health topics	99k
2	Mental Health Conversational Data *(Alexandreteles/Mental-health-conversational-data · Datasets at Hugging Face, 2023)*	Conversational data related to greetings.	661
3	Mental Health Counselling Conversations *(Amod/Mental_Health_Counseling_Conversations · Datasets at Hugging Face, 2001)*	Collection of questions on a wide range of mental health topics and answers provided by qualified psychologists.	3.5k
4	Mental Health Dataset *(Riyazmk/Mentalhealth Datasets at Hugging Face, n.d.)*	Scenario-based Question and Answers Dataset related to mental health	1.4k
5	Mental Health Chat Dataset *(Mpingale/Mental-health-chat-dataset · Datasets at Hugging Face, n.d.)*	Question Text and Answer Text columns selected as question and answer	2.7k
6	Psychology Dataset (J*khedri/Psychology-dataset · Datasets at Hugging Face, 2023)*	Chat questions and two responses were generated for each question asked	9.8k
7	SMILE Dataset (*Papers With Code - PsyQA Dataset, n.d.)*	Dataset for sentiment analysis in the context of mental health	54k
8	Psych8k Dataset *(EmoCareAI/Psych8k · Datasets at Hugging Face, n.d.)*	Dataset related to psychology	8k
9	Custom Chat Data	Question and answers formatted dataset specific to our Team and name of the chatbot	936

Technologies Used

Below are the state-of-art technologies used for the development of the project.

1. **Front-end**
    - **React.js:** A JavaScript library for building user interfaces.
    - **Tailwind CSS** *(Tailwind CSS - Rapidly Build Modern Websites Without Ever Leaving Your HTML., n.d.)*: It is a class with all the CSS properties written differently that can be directly used with simple shorthand.
    - **Particle.js** Garreau, V. (n.d.).**:** A lightweight JavaScript library for creating interactive particle animations on a webpage.
2. **Back-end**
    - **Fast API** (*FastAPI*, n.d.)- An efficient framework for building Python APIs. It is much faster than any other Python framework like Flask and Django. Fast API provides the flexibility to build large applications with high performance.
3. **Machine Learning**
    - **PyTorch:** A free and open-source library for machine learning and deep learning with more advanced built-in libraries to be utilized directly to ease the task of machine learning.
    - **Transformers:** The library enables developers to create sophisticated natural language processing (NLP) applications with minimal effort with pre-defined architecture.
4. **Database Technologies**
    - **MongoDB:** A powerful NoSQL database system that efficiently stores unstructured data.
    - **MongoDB Engine (MongoEngine):** A Python Object-Document Mapper (ODM) library for MongoDB, facilitating interaction with MongoDB databases in Python applications.
5. **Deployment Technologies**
    - **Atlas:** MongoDB Atlas is a MongoDB-created, managed, multi-cloud database service that simplifies deployment, management, and scaling, enabling developers to build resilient, high-performance applications across various cloud providers.
    - **AWS EC2:** Resizable, secure compute capacity is made possible by AWS EC2, which is perfect for implementing applications with scalable, on-demand infrastructure.

Methodologies

1. Literature Review

The process began with a basic literature review on some topics that seemed to be important yet unknown. This section of the report focuses on the topics which were needed to be studied that could be helpful in successfully building the chatbot. Some of them, along with appropriate brief research are explained in this section.

Transformers and Large Language Models: Transformers are very effective for training and inference since they handle input sequences in parallel. Compared to earlier recurrent neural network architectures like long short-term memory (LSTM), transformer models require less training time *(What Is a Transformer Model? | IBM, n.d.)***.** Large Language Models are built on transformer architectures and have received a great deal of attention and praise for their capacity to produce text that is coherent and relevant to context in a variety of fields. Pre-trained on copious amounts of text data, these models—like Google's BERT (Bidirectional Encoder Representations from Transformers) and OpenAI's GPT (Generative Pre-trained Transformer) series—learn rich representations of language patterns and semantics.

PyTorch: With PyTorch, Large Language Models (LLMs) and other machine learning projects can be effectively constructed. Because of its similarity to Python in design, it is straightforward to use and make experimenting and trying out new concepts a breeze (Alvi, 2024). It includes an extensive library of pre-trained models and tools, making it simple to get started and experiment with various approaches. Hence, it was decided that it will be well-suited for the project.

GPT (Generative Pre-trained Transformers): The GPT models are neural network-based language prediction models based on the Transformer architecture. Based on their language comprehension, they evaluate prompts, which are inquiries in normal language, and forecast the most appropriate response. The GPT models use the knowledge they acquire from training with hundreds of billions of parameters on large language datasets to accomplish that. They can produce more extended responses than just the following word in a sequence because they can dynamically attend to different input sections and consider the input context *(What Is GPT AI? - Generative Pre-Trained Transformers Explained - AWS, n.d.)*.

During each processing stage, the transformer neural network architecture employs self-attention mechanisms to focus on distinct portions of the input text.

Text Generation: To generate text that responds intelligently to user prompts or queries about mental health, it is necessary to take pertinent data and insights out of the text of these books. Using natural language processing methods such as keyword extraction, text summarization, and semantic analysis, chatbots can recognize important ideas, treatment approaches, and coping mechanisms highlighted in mental health literature. Chatbots may provide users seeking help with their mental health with invaluable advice, resources, and encouragement by utilizing the wealth of knowledge in mental health books. This helps users feel more connected and empowered as they work towards improving their mental health.

An Attempt at Implementation of Text Generation: Based on the research, the text generation model tried to create a chatbot that answers the user prompts from the knowledge fed into the GPT model. There were approximately 20 books fed to the model with different types of neural issues and their therapies with some deep medical terms and medications. The idea behind this chatbot was to chat with users to make them feel better and try some therapies at home, like exercise and meditation. Given the input text, the model only generated text and could not maintain a conversation, which does not align with the chatbot's use case. Moreover, the book data generated responses that suggest medicines for mentally disturbed people, which is unethical and against the doctor’s prescription, too cannot be sure of the correct medication indicated to the user, resulting in a dangerous condition.

Frontend and Backend Technologies: Research is required to determine the frontend and backend technologies to be used from the available options. At first, the idea was to research React, but there were options for backend technologies: Fast API and Flask. According to the research, Fast API is a newer and faster technology than Flask. Thus, it was decided to use Fast API and try to learn new technology.

2. Data Preprocessing

Several stages are involved in the data preprocessing pipeline to get the text data ready for additional modeling or analysis. These actions include removing named entities from the data to make it anonymous, translating the text to maintain linguistic consistency, adding greetings to improve the flow of the conversation, and combining the conversations into a single format to make it easier to perform tasks later, like sentiment analysis or chatbot model training. Every preparation stage helps to clean and standardize the data, which enhances the dataset's quality and usefulness for a range of natural language processing applications.

Named Entity Removal: The Distilbert-NER *(Sentence-transformers/Distilbert-base-nli-mean-tokens · Hugging Face, n.d.)* pre-trained model was used to identify entities like names of people, organizations, locations, dates, and other specific entities within the text data, which is the fine-tuned version of DistilBERT, which is a distilled variant of the BERT model. The dataset collected consisted of some names like Alex and Charlie. Once identified, remove the named entities from the text data to anonymize and protect sensitive information.

Text Translation from Chinese (SMILE Dataset) to English: Text translation converts text data from one language to another, facilitating language standardization and comprehension across different linguistic backgrounds. Google Translate API is a machine translation tool or library capable of translating text from Chinese to English. It ensures proper handling of linguistic nuances and idiomatic expressions during translation to preserve the meaning and context of the original text.

Enhance the Beginning of the Conversation with Greetings: Adding greetings at the beginning of conversations helps establish a polite and friendly tone, enhancing user engagement and conversational fluency. A dataset with specific greeting questions and a response was added to the beginning of each conversation and later fed to the model.

Merge Under Specific Format: Merging conversations under a specific format streamlines the text data for further analysis, modeling, or storage, improving accessibility and usability. Each row represents the conversation between human and gpt. The format of the dataset is human: <prompt> gpt: <response> human: <prompt> gpt: <response>.....

3. Data Exploration and Visualization

a. **Unigrams and Bigrams:** The chat data shows two horizontal bar charts displaying the most frequent words and word pairs. The first chart shows single words like "feel," "like," "time," "help," and "support," indicating frequent expressions of feelings and requests for assistance. The second chart shows word pairs like "feel like," "take time," "make feel," "self-esteem," and "self-care," indicating discussions often focus on personal experiences, emotional states, and self-improvement.

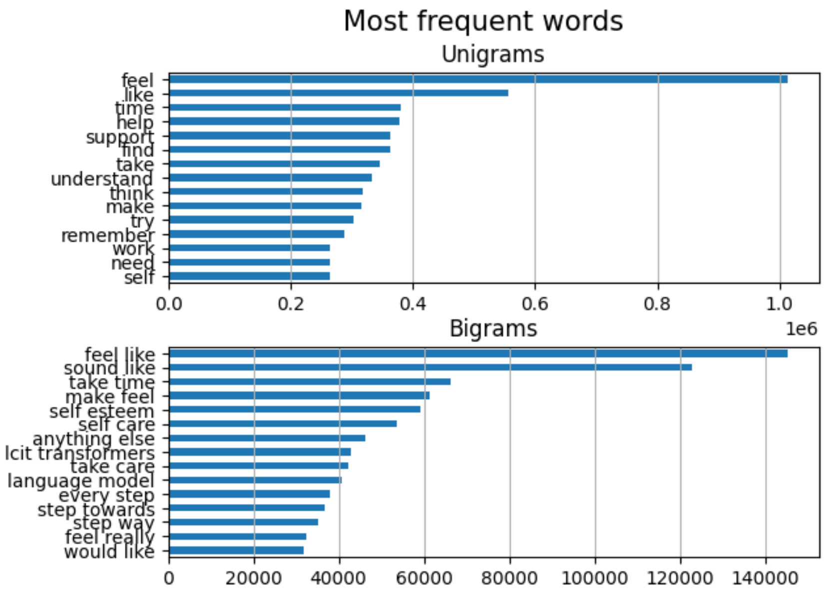

Figure 1 Unigrams and Bigrams of dataset.

b. **Word Cloud:** The word cloud below highlights the frequency of specific words like "self," "feel," "self-care," and "self-esteem" suggest a focus on personal well-being and emotional health. Smaller yet notable words like "support," "care," "understand," and "important" emphasize the supportive nature of the discussions.

Figure 2 Word Cloud of dataset.

c. **Latent Dirichlet Allocation (LDA):** The below chart represents the results of applying the LDA algorithm for topic modelling on a mental health conversational dataset. Each bubble corresponds to a specific topic identified by the LDA algorithm, with the size indicating the proportion of conversations or text passages associated with that topic.

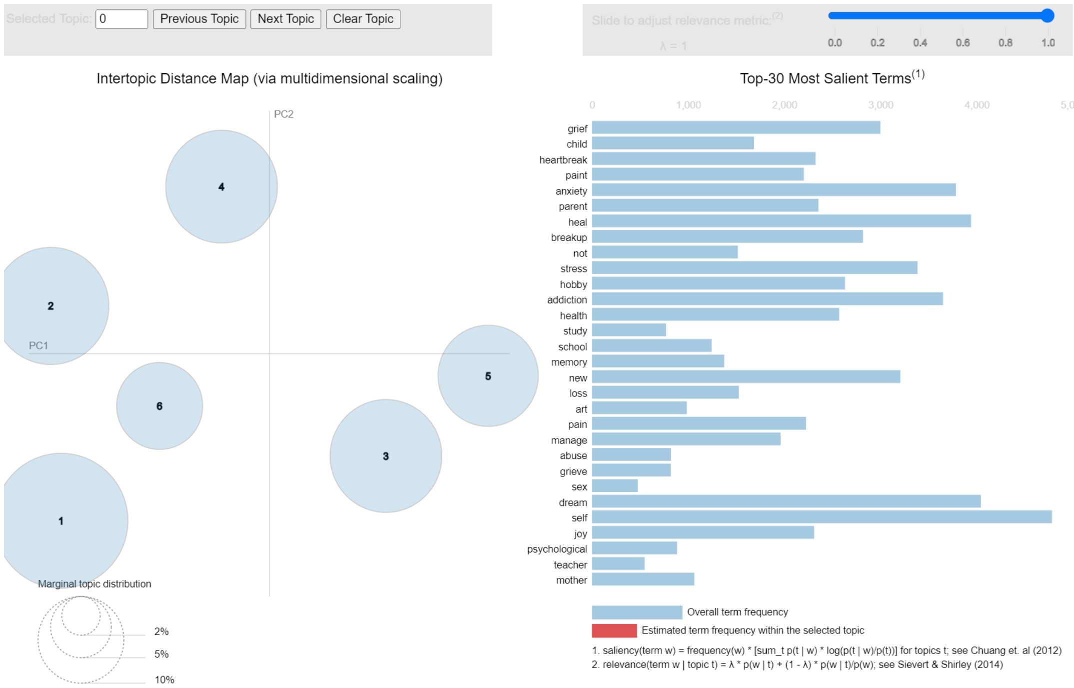

Figure 3 Latent Dirichlet Allocation on dataset.

d. Intent Classification: The below bar chart illustrates the results of intent classification categorizing each conversation into different intents or topics related to mental health issues. Each segment of the chart represents a specific intent, such as stress, anxiety, greetings, social anxiety, depression, and others, with the size of each segment indicating the proportion of conversations assigned to that intent.

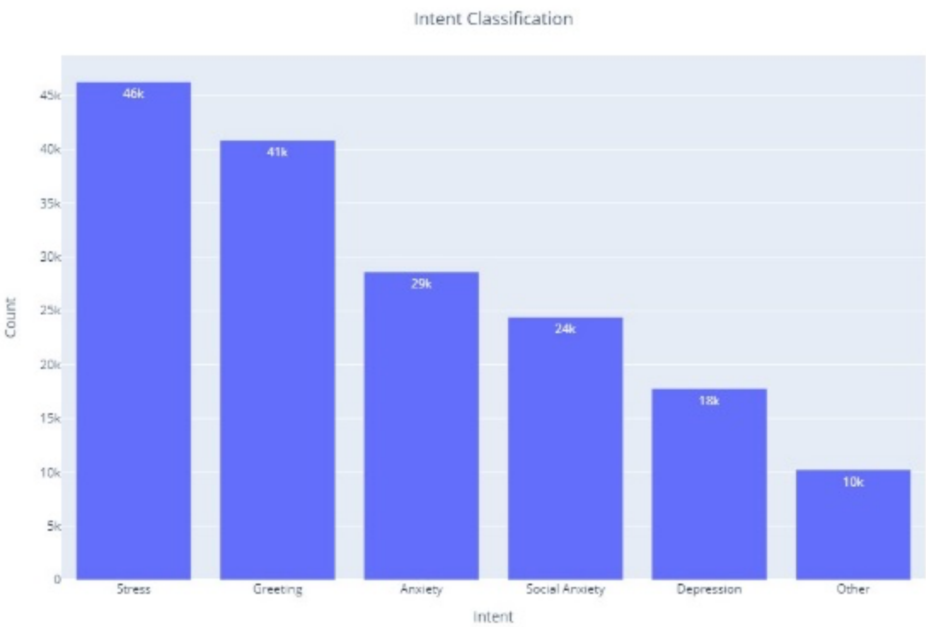

Figure 4 Intent Classification of conversations.

4. Methods Used

a. Tokenization

The GPTTokenizer, a Natural Language Processing (NLP) pipeline component, is essential for transforming unprocessed text input into a format the GPT model can comprehend and utilize efficiently. This procedure handles sophisticated language patterns like compound words and inflected forms and breaks the input text into meaningful units like words, punctuation, and special characters. To improve the model's comprehension and processing of sequential data, the GPTTokenizer also includes unique tokens, such as [CLS] (classification token) and [SEP] (separator token), to indicate the start and finish of input sequences. A pre-trained GPTTokenizer is used to tokenize the input data as it has a vast dictionary (*OpenAI GPT2*, n.d.).

b. Model Training

In developing our mental health chatbot, a pivotal stage was the training of the GPT-2 model to ensure its efficacy in delivering context-aware and empathetic responses to user inputs. However, to align its capabilities more closely with the nuanced requirements of mental health conversations, we undertook training from scratch with a similar configuration. This adaptation was made possible through the GPT-2 configuration (*OpenAI GPT2*, n.d.), a flexible framework that customizes the model's architecture and training dynamics. Key aspects such as the size of the model, the number of transformer layers, the learning rate, and training epochs were meticulously adjusted to cater to the specificities of the dataset and objectives.

The choice of GPT-2 as the underlying architecture was driven by its proven proficiency in generating coherent and contextually relevant text. By configuring the GPT-2 model's architecture and training parameters, aiming to balance responsiveness, and understanding, which is critical for addressing the complex nature of mental health inquiries. This approach enabled us to inject the model with a foundational knowledge uniquely suited to our objectives, ensuring it could engage users in meaningful, supportive dialogues. Training the model from scratch was a strategic choice, emphasizing the importance of constructing a chatbot that could navigate the intricacies of mental health support with empathy and insight. This foundational phase was crucial in creating a chatbot capable of positively impacting the mental health domain, highlighting the power of AI to offer genuine assistance. The model has 254M parameters.

c. Model Evaluation and Testing

Using a testing dataset generated with GPT4 LLM, the model's performance on unknown data was evaluated throughout the project's evaluation phase. Initially, all the data readily available was utilized for training since training LLM models requires a considerable amount of data to communicate with users and answer all possible queries effectively. However, more data was produced during the testing stage for different types of intents found in the training dataset to guarantee a thorough assessment. Specifically, 100 records were generated for every intent class, enabling a comprehensive evaluation of the model's performance and generalization skills on new and unseen data. This methodology allowed for a thorough assessment of the model's performance in managing real-world situations beyond the parameters of the training set. This also helped evaluate model performance on unbalanced data as per the data exploration.

1. **BLEU Score** *(Santhosh, 2023)**:*** A typical metric for assessing the quality of machine-generated text translations is the BLEU (Bilingual Evaluation Understudy) score, calculated by comparing the translations to human-generated references. It compares the generated text's n-gram precision—typically up to 4-grams—with the reference texts.
2. **ROUGE Score** *(Santhosh, 2023)**:*** By contrasting machine-generated summaries with human-generated references, the ROUGE (Recall-Oriented Understudy for Gisting Evaluation) score is a set of criteria to assess the quality of text summary and document generation. It assesses recollection of n-grams or up to 4 grams and how closely the generated summary and reference summaries match. The ROUGE score helps determine the coherence and content preservation of summaries.
3. **DistilBERT Score:** The DistilBERT (*Sentence-transformers/Distilbert-base-nli-mean-tokens · Hugging Face, n.d.)* architecture is a condensed form of the BERT (Bidirectional Encoder Representations from Transformers) model that serves as the foundation for a metric called the DistilBERT score. As a part of this project, the model used for computation is “sentence-transformers/distilbert-base-nli-mean-tokens.” It maps sentences and paragraphs to a 768-dimensional dense vector space and can be used for clustering or semantic search tasks. It evaluates the model's

efficacy or accuracy in a particular activity, like question-answering, text classification, or sentiment analysis.

Results

Some of the product results are shown below, including the front end, the working of the back end, and the model evaluation results.

1. **Landing Page:**

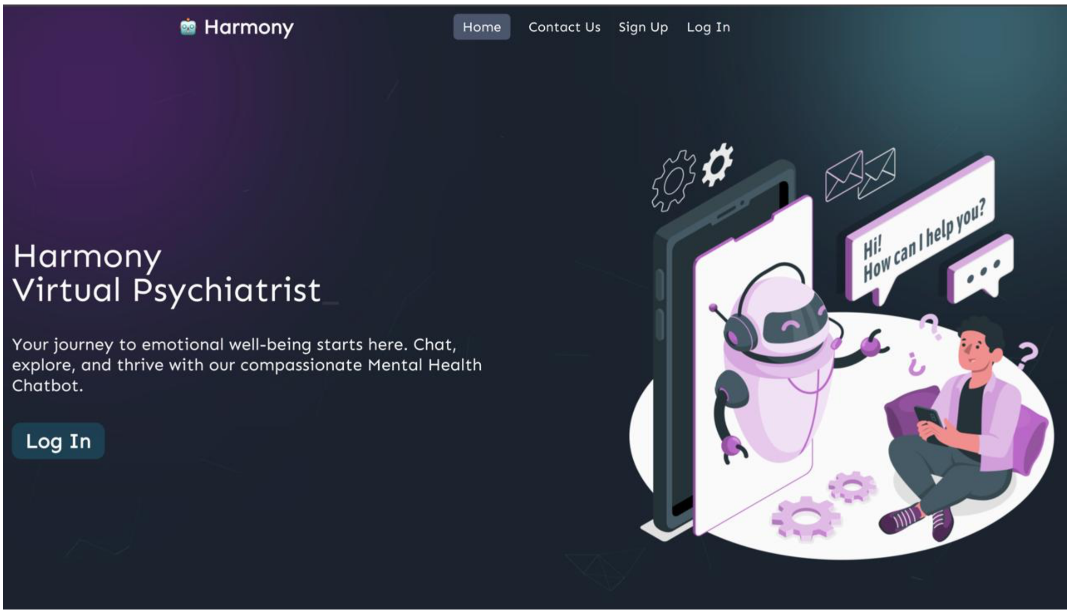

Figure 5 Landing Page of Webapp

2. **Sign Up Page:**

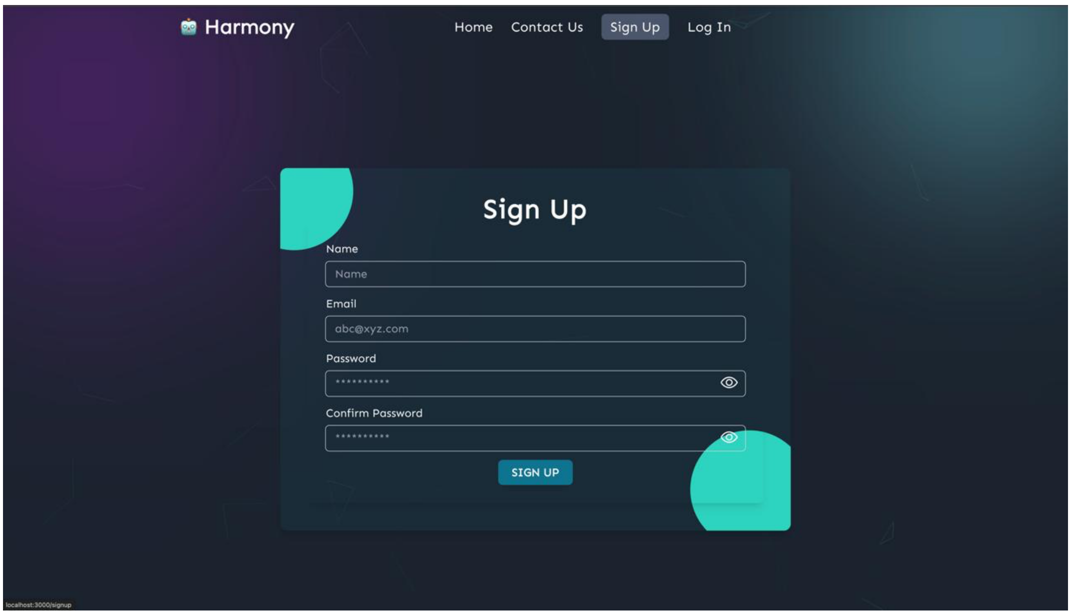

Figure 6 Sign Up Page

3. Login Page:

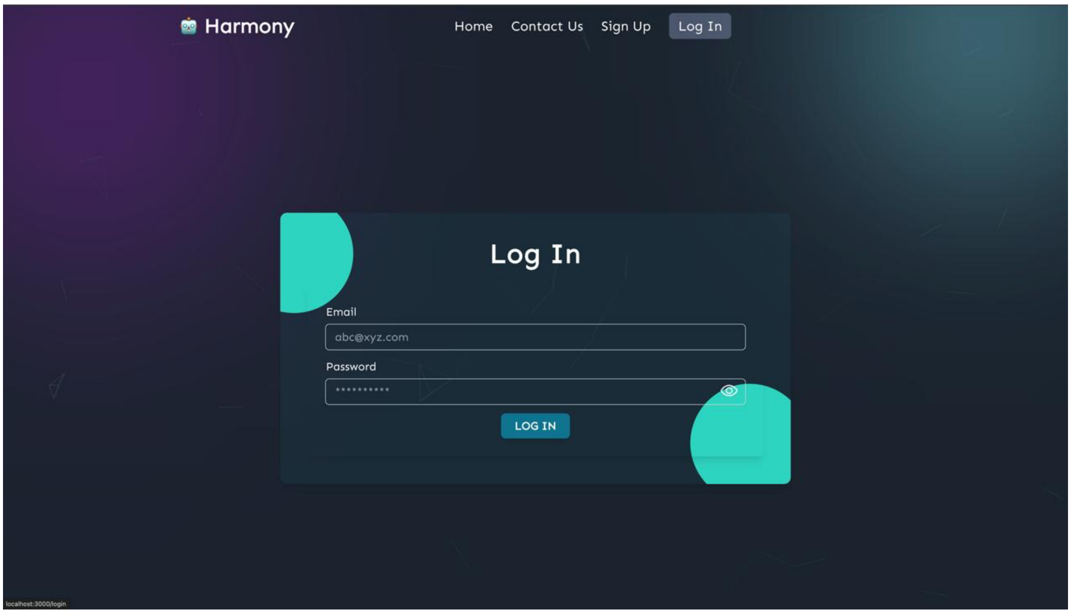

Figure 7 Login Page

4. Chat History:

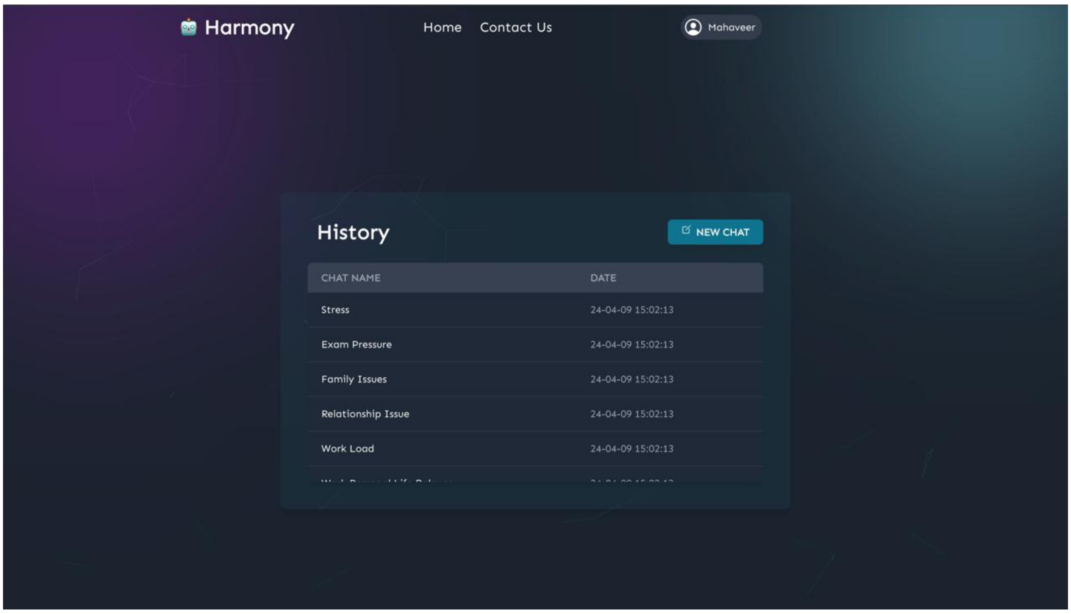

Figure 8 Chat History showing old chats with Harmony

5. Chat Screen and Conversation:

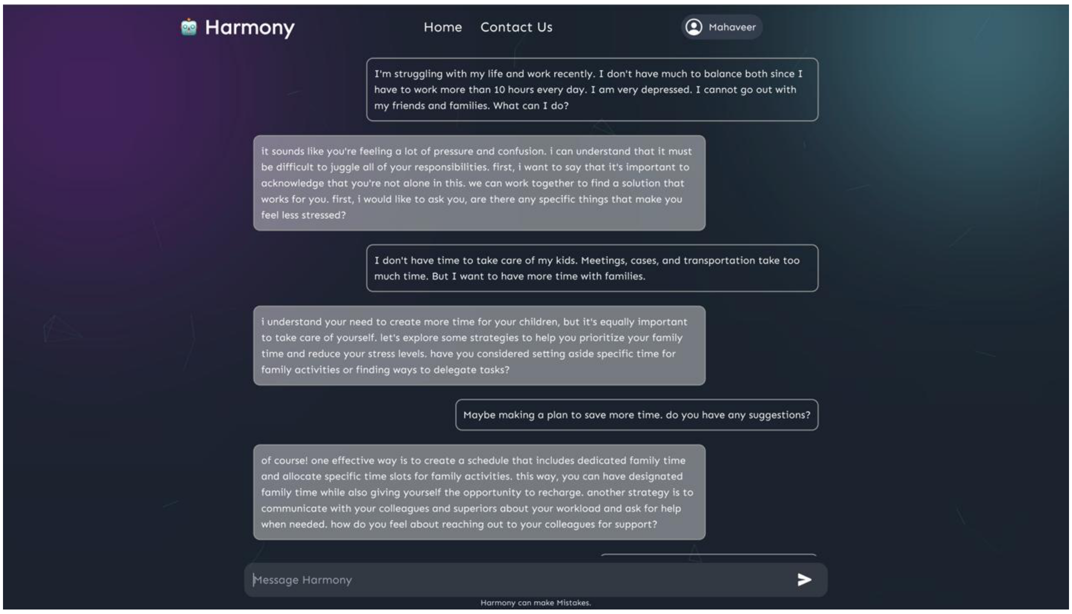

Figure 9 Chat Screen for user conversation

6. Evaluation Results:

	BLEU	ROUGE	Cosine Similarity
Stress	0.013695	0.173437	0.651334
Anxiety	0.017301	0.171054	0.611296
Depression	0.011467	0.160048	0.601141
Greeting	0.018578	0.096368	0.479845
Social Anxiety	0.009661	0.119167	0.596898
Others	0.013368	0.154787	0.507061

Figure 10 Evaluation Results using various metrics in the generated test data

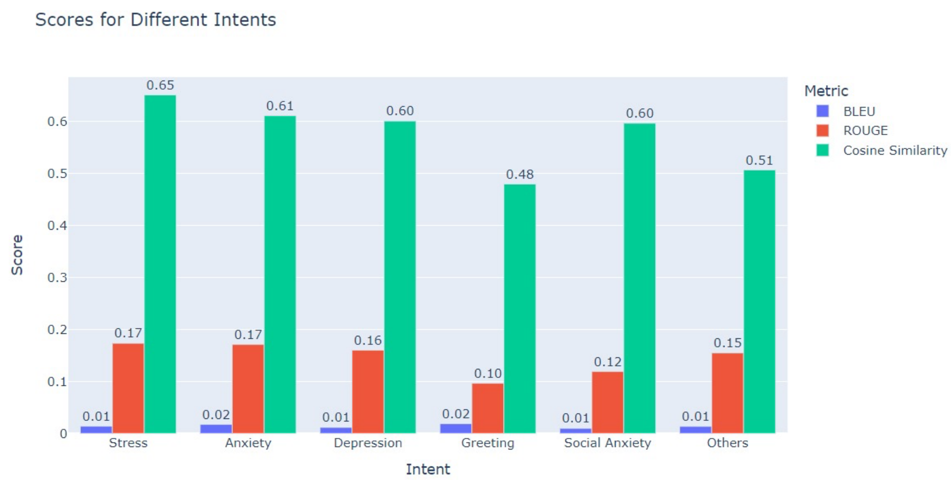

Figure 11 Evaluation Plot

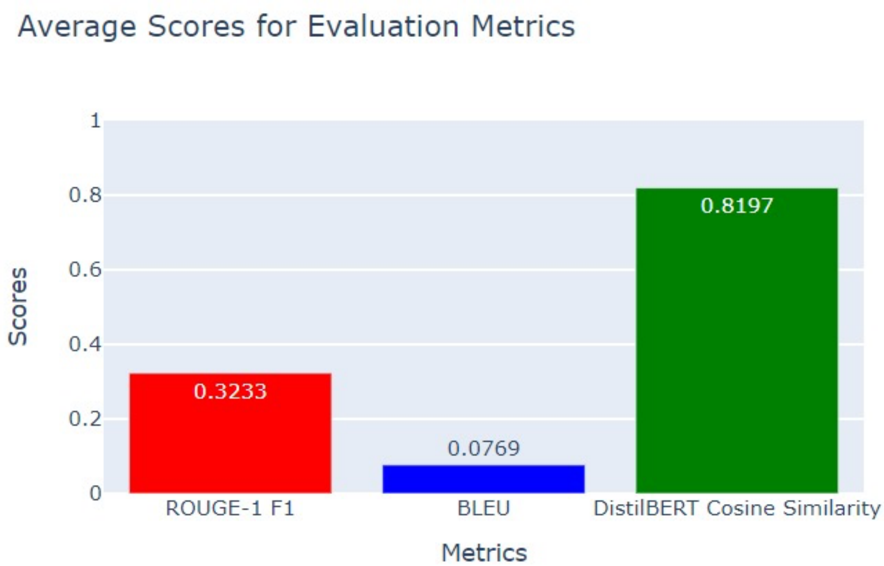

Figure 12 Evaluation on sample of Training Dataset

Discussion

According to the results shown in Figure 11, the BLEU score and ROUGE were evaluated to be around 0.1, which ideally should be more than 0.5. Thus, it is clear from the results that BLEU and ROUGE metrics could not evaluate the model response properly. However, the Cosine Similarity metrics gives better results on model evaluation. On the other hand, Figure 13 shows the evaluation of the chats chosen randomly from training dataset and compared generated responses with the actual responses. The performance of model with ROUGE score of 0.323 and BLEU score of 0.0769 indicates limited n-gram overlap with reference responses, which is common in open-domain conversations. Conversely, a Cosine

Similarity score of 0.8197 shows that the model does a good job of capturing the semantic core of the reference responses.

ROUGE and BLEU prioritize lexical similarity (n-gram overlap) however BLEU penalizes for outputs that are too brief. On the other hand, DistilBERT Cosine Similarity assesses semantic similarity, which might not be directly correlated with n-gram overlap, but nevertheless offers valuable information about how well the model captures intended meanings. Because legitimate responses in conversational models are inherently diverse, ROUGE and BLEU scores may be lower in these models.

Limitations

While mental health chatbots can provide valuable support and help, they may face limitations, as mentioned below:

1. **Medication Concerns:** It lacks the expertise to provide personalized medication advice, monitor medication side effects, or address complex medication-related questions. Users may also prefer interactions with healthcare professionals when discussing medication management to ensure safety and efficacy.
2. **Cross-Conversation Challenge:** One of the challenges this application faces is managing cross-conversations, where it is difficult to maintain context and coherence across the conversation while switching topics. This might lead to confusion and inefficiencies in understanding the context of the prompt while responding. Addressing cross-conversation issues requires sophisticated dialogue management algorithms and real-time context-tracking capabilities, which may pose technical challenges.
3. **Ending Conversations with "Bye":** While the mental health chatbot aims to provide continuous support, one specific shortcoming of this application is that the closing conversation statements, such as 'Bye,' are considered the end of the conversation.
4. **Limited Length of Conversations:** Mental health chatbots may face constraints in the length and depth of conversations due to technical limitations or preferences. The conversation length of the chatbot is only 1024 tokens. Thus, shorter conversations may limit the depth of exploration into users' concerns and restrict the effectiveness of the generated response.

Conclusion and Future Work

To conclude, the creation and assessment of the Virtual Psychiatrist chatbot mark a noteworthy advancement in the use of AI technology to promote mental health. The project demonstrates building a chatbot using some of the useful technologies such as, PyTorch, ReactJS, FastAPI, and MongoDB. The GPTTokenizer is used for tokenization while, GPT-2 model was trained from scratch that helped to reduce the project's constraints, which included the difficulty of capturing the complex vocabulary of mental health discourse. To guarantee the caliber of the training data, the project methodically carried out data extraction, cleaning, and visualization procedures with great care. GPTConfig parameters were utilized for training, highlighting the significance of customization for mental health applications. From various evaluation metrics available, BLEU, ROUGE, and DistilBERT Cosine Similarity measures were used for testing and evaluation, focusing on semantic relevance and context. The results emphasized the importance of using sophisticated assessment techniques like DistilBERT Cosine Similarity to capture the complex semantics of mental health discussions.

In the future, it could concentrate on several important topics to improve the functioning and impact of the mental health chatbot. Initially, to enhance the chatbot's comprehension and capacity to react to user inquiries accurately and sympathetically, continuously refining its training data and algorithms will be imperative. Secondly, integrating multimodal input and output modalities, including voice recognition, facial expression analysis, and gesture recognition, to enhance user engagement and accessibility can be another task that can be incorporated soon. This could create more immersive and engaging user experiences. By embracing innovative methodologies and continuous iteration, the Virtual Psychiatrist chatbot has the potential to evolve into a highly effective tool for providing empathetic and personalized mental health support, contributing to the well-being of individuals worldwide.

References

Garreau, V. (n.d.). *particles.js - A lightweight JavaScript library for creating particles*. particles.js - a Lightweight JavaScript Library for Creating Particles. https://vincentgarreau.com/particles.js/

FastAPI. (n.d.). https://fastapi.tiangolo.com/

Tailwind CSS - Rapidly build modern websites without ever leaving your HTML. (n.d.). Tailwind CSS. https://tailwindcss.com/

What is a Transformer Model? | IBM. (n.d.). Retrieved March 31, 2024, from https://www.ibm.com/topics/transformer-model

Alvi, F. (2024, January 24). PyTorch vs TensorFlow in 2024: A Comparative Guide of AI Frameworks. OpenCV. Retrieved March 31, 2024, from https://opencv.org/blog/pytorch-vs-tensorflow/#:~:text=PyTorch%20vs%20TensorFlow%3A%20Both%20are,for%20beginners%20and%20rapid%20prototyping.

Santhosh, S. (2023, April 17). *Understanding BLEU and ROUGE score for NLP evaluation*. Medium. https://medium.com/@sthanikamsanthosh1994/understanding-bleu-and-rouge-score-for-nlp-evaluation-1ab334ecadcb

sentence-transformers/distilbert-base-nli-mean-tokens · Hugging Face. (n.d.). https://huggingface.co/sentence-transformers/distilbert-base-nli-mean-tokens

dslim/distilbert-NER · Hugging Face. (2001, January 18). https://huggingface.co/dslim/distilbert-NER

jerryjalapeno/nart-100k-synthetic · Datasets at Hugging Face. (2001, February 20). https://huggingface.co/datasets/jerryjalapeno/nart-100k-synthetic

alexandreteles/mental-health-conversational-data · Datasets at Hugging Face. (2023, September 27). https://huggingface.co/datasets/alexandreteles/mental-health-conversational-data

Amod/mental_health_counseling_conversations · Datasets at Hugging Face. (2001, March 4). https://huggingface.co/datasets/Amod/mental_health_counseling_conversations

Riyazmk/mentalhealth · Datasets at Hugging Face. (n.d.). https://huggingface.co/datasets/Riyazmk/mentalhealth

mpingale/mental-health-chat-dataset · Datasets at Hugging Face. (n.d.). https://huggingface.co/datasets/mpingale/mental-health-chat-dataset

jkhedri/psychology-dataset · Datasets at Hugging Face. (2023, October 6). https://huggingface.co/datasets/jkhedri/psychology-dataset

Papers with Code - PsyQA Dataset. (n.d.). https://paperswithcode.com/dataset/psyqa

EmoCareAI/Psych8k · Datasets at Hugging Face. (n.d.). https://huggingface.co/datasets/EmoCareAI/Psych8k