

Title: Leveraging AI Chatbot Technology for Mental Health Support

Author(s): Karen Dang & Sharief Nur

Abstract:

Mental health disorders affect millions of individuals worldwide, with significant implications for well-being and quality of life. However, access to timely and effective mental health support remains a challenge due to various barriers, including stigma, limited resources, and geographical constraints. This project proposal aims to address these challenges by leveraging artificial intelligence (AI) chatbot technology to provide accessible and scalable mental health support interventions.

Objectives:

The project aims to develop an AI chatbot platform specifically designed to cater to the requirements of individuals seeking mental health support. This platform will offer evidence-based interventions, psychoeducation, and crisis support through its user-friendly interface. A key feature of the platform will be its ability to personalize interventions according to user preferences, needs, and the context of the conversation. To ensure the platform's efficacy and user satisfaction, rigorous testing will be conducted, and feedback from users will be actively sought to evaluate its effectiveness, usability, and overall performance.

Methods:

This project will utilize the capabilities of Generative Pre-trained Transformer (GPT) models, renowned for their ability to generate coherent and contextually relevant text based on input sequences. GPT models leverage self-attention mechanisms to capture contextual information from input sequences, thereby enabling them to process and generate text with a high degree of accuracy and fluency.

The proposed GPT-based mental health chatbot will capitalize on the applications of GPT in mental health support. Researchers have explored the potential of GPT-based chatbots in various domains, including symptom monitoring, psychoeducation delivery, crisis intervention, and therapy augmentation. GPT's unique ability to comprehend context and nuances in language enables chatbots to provide personalized and empathetic responses, thus enhancing user engagement and satisfaction. Studies have demonstrated the feasibility and effectiveness

of GPT-based chatbots in engaging users, providing personalized support, and delivering evidence-based interventions tailored to individual needs.

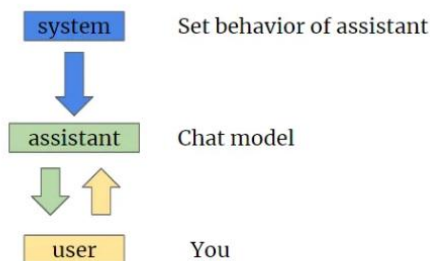
GPT and Prompt Engineering:

Generative Pre-trained Transformer (GPT) models, developed by OpenAI, are at the forefront of natural language processing (NLP) technology. These models are renowned for their ability to generate coherent and contextually relevant text based on input prompts. Prompt engineering plays a crucial role in harnessing the full potential of GPT models. By carefully crafting prompts tailored to specific tasks or contexts, users can guide the model to generate responses that align with their objectives. Effective prompt engineering involves providing clear instructions and context to the model, enabling it to produce accurate and relevant outputs. This approach empowers users to utilize GPT models for a wide range of applications, from text generation and summarization to translation and dialogue systems.

Training GPT Models on Specific Prompts:

Training a GPT model on certain prompts involves fine-tuning the pre-trained model using task-specific data and prompts. The process typically begins by collecting a dataset relevant to the desired task or domain. Next, researchers or developers preprocess the data and design prompts that encapsulate the desired inputs and outputs. These prompts serve as guiding cues for the model during training, shaping its understanding of the task and desired outcomes. The fine-tuning process involves feeding the model with prompt-response pairs and adjusting its parameters to optimize performance on the target task.

Through iterative training and evaluation, developers refine the model's ability to generate accurate and contextually appropriate responses to the specified prompts. This approach enables users to tailor GPT models to their specific needs and domains, unlocking their full potential for diverse applications in natural language understanding and generation.

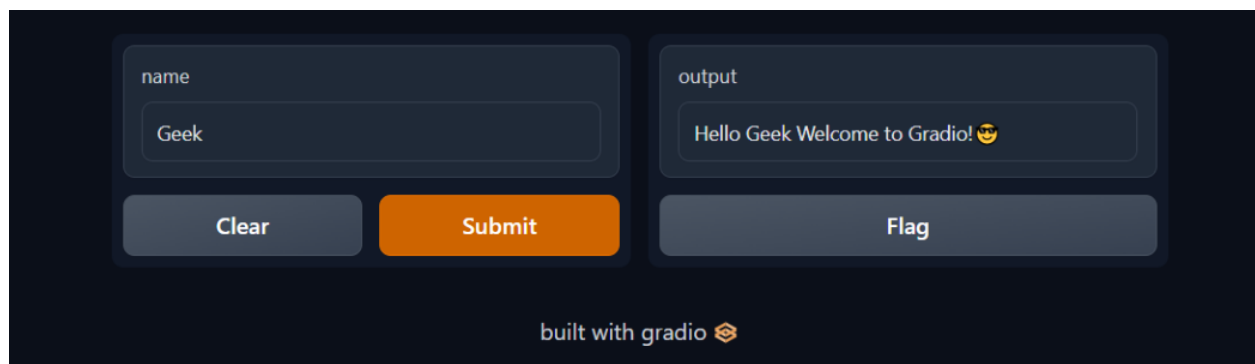


Leveraging AI Chatbot Technology for Mental Health Support

After training the GPT model with prompt engineering techniques tailored to mental health conversations, integrating it with Gradio for deployment enhances accessibility and usability. Gradio simplifies the process of creating interactive interfaces, allowing users to interact with the GPT model seamlessly.

Through Gradio's customizable components such as text boxes and sliders, users can input their mental health queries and receive empathetic and supportive responses from the GPT model. This integration bridges the gap between advanced AI models and end-users, making mental health support more accessible and user-friendly.

With Gradio, deploying the prompt-engineered GPT model becomes a straightforward process, empowering developers to bring their AI-driven mental health solutions to a wider audience with ease.



Project Implementation:

Data Collection: Gather conversational data relevant to mental health support.

Prompt Engineering: Design prompts to guide the chatbot's responses and ensure they are contextually relevant and empathetic.

Model Training: Fine-tune the GPT model using the collected data and prompts to enhance its ability to generate appropriate responses in mental health-related conversations.

Gradio Interface Development: Create a user-friendly interface using Gradio to facilitate interaction between users and the chatbot.

Testing and Evaluation: Conduct rigorous testing and evaluation to assess the chatbot's performance in providing accurate, helpful, and empathetic responses.

Iterative Improvement: Incorporate user feedback and iteratively improve the chatbot's performance, usability, and effectiveness.

Leveraging AI Chatbot Technology for Mental Health Support

The following examples are representative of various Training Data and the associated Output:

Training_data:

```
# Define user prompts and bot responses for training
training_data = [
    {"prompt": "I'm feeling anxious and stressed.", "response": "It's okay to feel that way. Have you tried relaxation techniques?"},
    {"prompt": "I'm feeling depressed and hopeless.", "response": "I'm here to support you. Consider talking to a therapist."},
    {"prompt": "I'm struggling with my self-esteem.", "response": "Remember you are worthy. Focus on your strengths."},
    {"prompt": "I'm having trouble sleeping.", "response": "Try establishing a bedtime routine and avoiding screens before bed."},
    {"prompt": "I'm feeling overwhelmed with everything going on in my life.", "response": "Break tasks into smaller steps and ask for help."},
]
```

Output:

The image displays three screenshots of the 'Mental Health Support Chatbot' interface, each showing a user prompt and the chatbot's response. The interface is dark-themed with a white header and footer. The chat area is a rounded rectangle with a light gray background. Below the chat area are two buttons: 'Clear' (gray) and 'Submit' (orange). Below the chat area is a 'Response' section with a light gray background and a 'Flag' button (gray) below it.

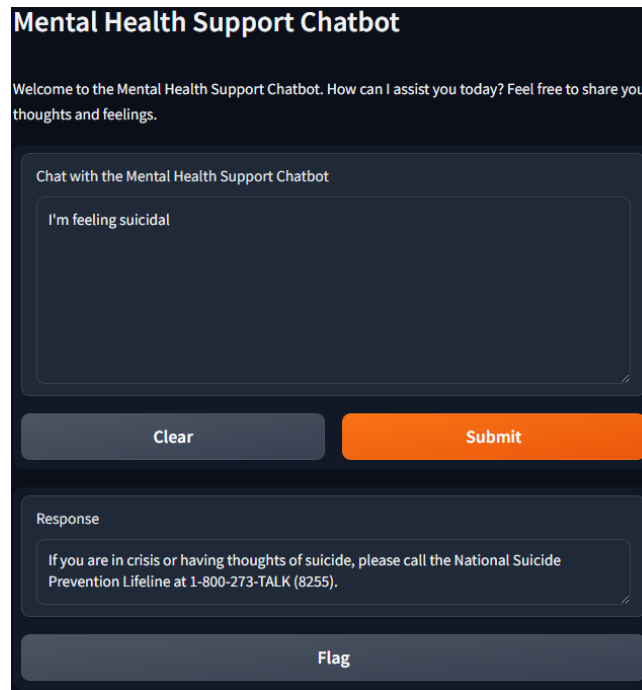
Screenshot 1: User prompt: "I'm feeling depressed and hopeless". Response: "I'm here to support you. Consider talking to a therapist."

Screenshot 2: User prompt: "I'm feeling anxious and stressed". Response: "It's okay to feel that way. Have you tried relaxation techniques?"

Screenshot 3: User prompt: "I'm having trouble sleeping". Response: "Try establishing a bedtime routine and avoiding screens before bed."

Looking for key_words “suicidal” to prompt for defined response:

```
mental_health_keywords = {
    "anxious": "mild",
    "depressed": "moderate",
    "stressed": "mild",
    "self-esteem": "mild",
    "overwhelmed": "moderate",
    "sleeping": "moderate",
    "suicidal": "severe"
}
input_text_lower = input_text.lower()
for keyword, severity in mental_health_keywords.items():
    if keyword.lower() in input_text_lower:
        if keyword.lower() == "suicidal":
            return "severe", "If you are in crisis or having thoughts of suicide, please call the National Suicide Prevention Lifeline at 1-800-273-TALK (8255)."
```



Conclusion:

In summary, the utilization of prompt engineering techniques to refine the GPT model has significantly improved the chatbot's responsiveness and accuracy. Through fine-tuning the model with specific prompts, we have empowered the chatbot to accurately interpret and comprehend the provided prompts, as well as identify predefined keywords within the user inputs. This advanced capability enables the chatbot to deliver customized responses that align with the user's intentions, even when the input slightly deviates from the original prompt. Through this method, we have effectively given the chatbot the capacity to offer precise and relevant assistance, thereby enhancing both the user experience and the effectiveness of the mental health support platform.

Resources:

<https://blog.hubspot.com/website/python-ai-chat-bot>

<https://github.com/deeepss/SpacyBot>

<https://www.linkedin.com/pulse/basic-chatbot-using-spacy-library-pradeepa-loganathan/>

https://medium.com/@pankaj_pandey/spacy-llm-integrating-llms-into-structured-nlp-pipelines-7134dd05ebc2

<https://spacy.io/usage/large-language-models>

<https://francisgichere.medium.com/sentiment-analysis-of-app-reviews-a-comparison-of-bert-spacy-textblob-and-nltk-9016054d54dc#:~:text=If%20you%20are%20looking%20for,spaCy%20is%20a%20better%20choice.>

https://medium.com/@Ben_Obe/introduction-to-nlp-transformers-and-spacy-8ac9539f3bc1

<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC10242473/>

<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9643933/>

<https://www.npr.org/sections/health-shots/2023/01/19/1147081115/therapy-by-chatbot-the-promise-and-challenges-in-using-ai-for-mental-health>

<https://www.geeksforgeeks.org/prompt-engineering-for-chatbot/>

<https://medium.com/@flavioitoriano/creating-and-improving-a-chatbot-with-prompt-engineering-b39d2a0c6116>

<https://medium.com/@tariqsaad1997/chatgpt-prompt-engineering-part-4-building-a-customized-chatbot-165db7515c29>