Digital Counsellor:
Generative based chatbot using a rank based approach on training data

Manav Saini, Navidha Jain, Tanishqa Shital Singh, Anishka, Ashwin Tomer, Dux Pal Singh

1 Introduction

In recent years, the use of digital technology has transformed the field of mental health care. One area in which this technology has been particularly useful is in the development of digital counselling chatbots. These chatbots are designed to provide support and guidance to people experiencing mental health issues by engaging in conversation with them. As the world becomes increasingly digital, it is important to explore how technology can be used to help those in need of mental health support. Mental health issues are a growing concern worldwide, and the demand for mental health services is increasing. However, many people are hesitant to seek help due to stigma, lack of access to services, or concerns about the cost of care. In addition, mental health professionals may not always be available when people need support.

Digital counselling chatbots have the potential to address some of these challenges by providing accessible and affordable mental health support to a wide range of people. This project aims to develop a digital counselling chatbot that is designed to offer support and guidance to people experiencing mental health issues. The chatbot will be designed to provide a safe and confidential space for users to discuss their concerns and receive practical advice and information.

2 Problem Statement

Our goal is to develop a chatbot using a generative model that uses a rank-based approach training data. To obtain this training data, we will scrape data from Counsel Chat, a platform that hosts certified counselors whose answers are ranked with upvotes. We will use this data with GPT2, a generative model that has demonstrated state-of-the-art performance in tasks such as question answering, summarization, reading comprehension, and translation[4].

3 Background

Mental health issues are a major concern worldwide. These disorders can have a significant impact on a person's quality of life and can also affect their ability to work, socialize, and carry out daily activities. While mental health services are available in many countries, there are still significant barriers to accessing these services, including stigma, lack of availability, and high costs.

The use of digital technology in mental health care has been shown to have a positive impact on outcomes, and there is a growing body of evidence to support the use of digital interventions in mental health care. Chatbots are computer programs that are designed to simulate conversation with human users, and they can be programmed to provide support and guidance to people experiencing mental health issues. In particular, digital counselling chatbots have been shown to be effective in providing support and guidance to people experiencing mental health issues. They offer a potential solution to these challenges by providing accessible and affordable mental health support. These chatbots are available 24/7 and can be accessed from anywhere with an internet connection, making them a convenient and accessible option for people in need of support.

4 Related Works

[1]Chatbot for Mental Well-being: This paper aims to develop a therapy chatbot that can provide easy access to hassle-free mental health services. The proposed chatbot uses two major modules: an SVM classifier to detect the user's mood based on their input and a Seq2Seq model that generates appropriate responses to the user's input. From the Dataset each of the 32 Emotions was mapped to a unique number. The SVM algorithm generates a numerical output that corresponds to a particular mood based on the input provided by the user, while the Seq2Seq model uses an encoder-decoder architecture to generate an appropriate response to the user's input. The project utilizes the Empathetic Dialogues dataset for mood classification and the Keras and Sklearn libraries for model development. The chatbot's primary goal is to help ease the user's state of mind and keep track of their mood over a span of time.

[2]A Chatbot for Psychiatric Counseling in Mental Healthcare Service Based on Emotional Dialogue Analysis and Sentence Generation: The paper discusses the use of artificial intelligence (AI) methods to recognize human emotions and provide conversational counseling services for mental health care using a counseling chatbot. The chatbot uses deep learning-based emotion classification models

and natural language processing (NLP) methods to analyze consultation content and provide appropriate responses. The paper focuses on emotion recognition and monitoring, conversation understanding on the chat assistant, and developing a personalized dialog system that communicates emotionally with the user. The chatbot also uses additional user information such as facial expression, age, sex, spatial context, location context, and bio-signals collected via wearable devices. The goal is to develop a personalized, user-customized correspondence technology that communicates with users through speech, text, audio, and visual representation based on the user's age, gender, and recognized emotions. The paper also discusses the importance of tracking persistent emotional changes for improving the effectiveness of counseling. The chatbot uses a method that is a user-customized correspondence technology that communicates with users through speechtext-audio-visual representation based on the user's age gender classification and recognized emotions.

[3]Counsellor Chatbot: The article presents a study on developing a chatbot named Xen, which uses Retrieval and Generative techniques, including AIML and K-Means self-learning, to provide counseling services and advice based on user input. The chatbot is trained on three datasets, resulting in three types of chatbots: Retrieval Pattern Matching, Retrieval Rule-Based AIML, and Generative. The text describes the different chatbot architectures, including retrieval-based, AIML-based, and generative, and also discusses techniques such as NER and skip-thought vectors for improving chatbot performance. Overall, the article provides an overview of various techniques used in chatbot development.

[4]An Evaluation of Generative Pre-Training Model-Based Therapy Chatbot for Caregivers: This study describes the development and evaluation of a chatbot designed to provide mental health support to caregivers. The chatbot was developed using the GPT-2 generative pre-training model, which was fine-tuned on 306 therapy session transcripts involving 152 family caregivers. Three evaluation criteria were used, namely the proportion of non-words generated, sentence length, and sentiment analysis. The findings showed that the fine-tuned model generated a higher proportion of non-words, but performed better in terms of sentence length and sentiment analysis when compared to the GPT-2 model. The responses generated by the fine-tuned model closely resembled those of an actual therapist, in contrast to the nonsensical responses generated by the GPT-2 model. The sentiment in the fine-tuned model's responses was more positive than in the GPT-2 model, but also had more negative sentiment responses than in actual therapists' responses. The study highlights the limitations of generative-based chatbots for therapy contexts, such as safety concerns, credibility, personality suitability, and nuanced responses. The findings also suggest that designing human-AI interaction for therapy contexts is challenging due to unpredictable responses and the need for larger training datasets. Despite these limitations, the fine-tuned model has the potential to generate therapist-like responses that can provide support to caregivers, especially in situations where access to actual therapists may be limited or difficult. Further research is needed to improve the accuracy and effectiveness of chatbots for mental health support.

[5]The article highlights the growing demand for mental health services and the potential of chatbots to fill the gap. The chatbot is designed to provide support to users in a conversational way, and can offer customized advice and resources based on the user's responses. It outlines the process of developing the chatbot, including designing the user interface, creating a database of responses, and testing the bot with users. The article discusses the development of a demo chatbot that aims to make accessing mental health information more interactive and engaging for users. The chatbot uses emojis to allow users to select their mood or feeling, and then provides information and tips on various mental health issues. The article also suggests that chatbots could be used as a more user-friendly way of screening for mental health issues. The chatbot is designed to be incorporated into a wider web application, and the article discusses the use of AI and ethical considerations in the development process.The article highlights the potential of chatbots to provide accessible and personalized mental health support, but emphasizes the importance of ongoing development and evaluation to ensure their effectiveness and safety.

[6] The research paper "Unsupervised Summarization of Psychotherapy Data" by Aaron Z. Reed, from the Department of Physics at Stanford University, explores the potential of using unsupervised text summarization techniques to summarize psychotherapy session transcripts. The study is motivated by the shortage of psychotherapists in the US and the need for effective therapy chatbots that can provide valuable data to therapists and reduce their workload. The study focuses on two unsupervised text summarization methods: extractive and abstractive. The extractive approach selects the most relevant sentences from the original text to create a summary, while the abstractive approach generates a summary that may contain new phrases and sentences not present in the original text. The dataset used in the study is compiled from Counsel Chat, a website where users can ask questions that are answered by qualified therapists. The dataset contains around 2,000 questions with their responses provided as a CSV file, but there are no labels or summaries included with the dataset.

Therefore, the learning task was reframed in terms of unsupervised text summarization. The authors used the SkipThoughts model as a baseline for the extractive summarization approach. SkipThoughts is an encoder-decoder model that encodes sentences into vectors and predicts surrounding sentences using skip-thought vectors. The authors also compared the results and training times of QuickThoughts, an elaborated version of the SkipThoughts model. For the abstractive summarization approach, the authors used the MeanSum algorithm, which uses an attention mechanism to assign weights to each sentence in the original text based on its relevance to the summary. The model then generates a summary by selecting the most relevant sentences and generating new sentences using a language model. The study compared the results of both approaches subjectively and with automated metrics. The authors found that both the extractive and abstractive approaches are promising for automated summarization in the context of psychotherapy. The abstractive approach was found to be more effective in capturing the general consensus among responses, while the extractive approach produced more concise summaries. Overall, the study demonstrates the potential of unsupervised text summarization techniques in the context of psychotherapy, and highlights the need for high-quality therapy chatbots that can provide valuable data to therapists and improve the quality of care provided to clients.

[7] The text discusses the use of mobile applications, chatbots, and social robots for delivering psychosocial interventions to assist mental health practitioners in reducing their workload. The text provides examples of several chatbots, including EMMA and SERMO, which use Ecological momentary interventions and Cognitive Behavioural Therapy respectively. The article also mentions the Mental Health Data Set, NHS Digital, and Counsel Chat as resources for mental health professionals. The text also explains how Convolutional Neural Networks (CNN), Recurrent Neural Networks (RNN), and Hierarchical Attention Networks (HAN) are effective in carrying out tasks such as tokenizations, emotion detection, sentence classification, text classification, sentiment analysis, text summarization, and machine translation, and how these networks are used in the development of chatbots. The evaluation metric were words per session as a person suffering through any mental health disorder won't choose to talk to a chat bot about their issues and sentiment analysis which can have multiple levels, most commonly they are divided in five, very negative, negative, neutral, positive and very positive. Emotion analysis is the technique used to analyse the mood of the user, in terms of percentage positivity or negativity. First all sentences are analysed and assigned a mood. Then using an algorithm, percentage positivity or negativity is calculated. The text concludes by discussing a model which is transformer based that identifies mental states based on emotional labels.

[8] The paper "Chatbot for Mental Well-being" proposed a project that aims to develop a generative chatbot to help people relieve stress by providing a medium to talk to. The chatbot will use two major modules: an SVM classifier to detect the user's mood and a Seq2Seq model to generate an appropriate response. For mood classification, the Empathetic Dialogues dataset, containing 12,424 conversations grounded in emotional situations with 32 emotion labels, was used. For the Seq2Seq model, the Counselchat dataset, with 1659 question-answer pairs from multiple therapists, was used. The data preprocessing involved mapping each emotion label to a unique number, using regular expressions to remove special symbols, punctuation, and links, removing stop words, and performing word-level tokenization. A Term Frequency-Inverse Document Frequency vectorizer and a RandomOverSampler were used to balance the data. The Seq2Seq model was developed using Keras and incorporated an additive attention mechanism proposed by Bahdanau et al. GloVe Embedding was used to create the embedding layer. The results showed that the SVM classifier achieved an accuracy of 86.64% on the test dataset, indicating that it is effective in detecting the user's mood. The Seq2Seq model achieved a BLEU score of 0.61, indicating that the generated responses were comparable to those of human therapists. The model also showed promising results in handling out-of-vocabulary words and achieving semantic coherence in the generated responses. Overall, the proposed project demonstrates the potential of using machine learning techniques to develop a chatbot that can effectively communicate with users and help them relieve stress. The use of the Empathetic Dialogues and Counselchat datasets, along with appropriate preprocessing techniques and model architecture, resulted in a highly accurate and effective chatbot. Further improvements can be made by incorporating additional datasets and fine-tuning the model to better handle complex emotions and responses.

5 Objectives and Goals

Our idea is to develop a digital counseling chatbot that utilizes a combination of NLP algorithms, Machine Learning (ML) techniques, and personalized counseling sessions. The retreival based design uses a predefined set of queries and responses that are fed into the system and uses a decision making tree which makes is insufficient for multilinear conversations. It restrains free conversations and will fail the task if users' inputs do not match any database, making it difficult to improve usability. The rule-based chatbots are structured as a dialog tree and often use regular expressions to match a user's input to human-like responses. Here, only the previous output is taken into consideration to respond in relation to the context. We chose a generative model

based design as they allow for conversational flexibility. This model applies machine learning techniques to train the chatbots to learn and generate responses based on a large amount of training data. This makes them more intelligent as they take word by word from the query and generates the answers. The chatbot can also be fine-tuned with different domain data for unique purposes for its target users.
No metric can define whether a particular chatbot is good or effective because of the lack of human judgement while testing using metrics like BLEU, METEOR, and ROUGE-N. Therefore we use human judgement criteria before training the model in the form of upvotes. Here, the upvotes play an important role as it gives us a metric of how appropriate those responses are according to human judgement. This data can be used with the profile data of counsellors to create a high quality dataset needed for generative models to provide a solution for digital counselling and this is where our novelty lies.

6 Techniques/Algorithms

Ranking based approach for a high quality dataset:
1)Upvotes (human judgement metric)
2)Profile data of Counsellors
For the generative model we have used the medium version of GPT2 and have finetuned it on the dataset created. GPT2 was chosen because it has achieved state-of-the-art performance on language tasks like question answering, reading comprehension, summarization, and translation[4].

7 Baseline Evaluation and Metrics

For the baseline model we have directly used the dataset which is available by Counsel Chat website. We have used BLEU, METEOR and NIST metrics to evaluate the technical performance of our baseline model. The following scores were obtained:

METRIC	SCORE
bleu_2	0.4451666503993174
bleu_4	0.2939982596113935
meteor	0.3861091928395942
nist_2	3.053938345854437
nist_4	3.4386304341956144

Table 1: Baseline Model Performance

8 Potential Contributions

Dux Pal Singh, Anishka: Data collection, pre-processing of data, identification of different models to be tested, final evaluation
Manav Saini, Navidha Jain: Implementation of different models for generation task and choosing the best one, Improving baseline scores
Ashwin Tomar, Tanishqa Shital Singh: Implementation of different models for response generation and choosing the best one, Improving baseline scores
Report writing: All members

9 Citations and Bibliographies

[1] https://www.itm-conferences.org/articles/itmconf/abs/2021/05/itmconf_icacc2021_03019/itmconf_icacc2021_03019.html

[2] https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=7962482&tag=1

[3] https://www.academia.edu/download/56261875/07.MRCS10087.pdf

[4] https://arxiv.org/pdf/2107.13115.pdf

[5] https://www.scienceopen.com/hosted-document?doi=10.14236/ewic/HCI2017.24

[6] https://ieeexplore.ieee.org/document/9824269

[7] http://cs230.stanford.edu/projects_fall_2020/reports/58656561.pdf

[8] https://www.itm-conferences.org/articles/itmconf/abs/2021/05/itmconf_icacc2021_03019/itmconf_icacc2021_03019.html