Generative Interactive Psychotherapy Expert (GIPE) Bot

Ayesheh Ahrari Khalaf [1†], Prof. Dr. Aisha Hassan Abdalla Hashim[2††], Asst. Prof. Dr. Akeem Olowolayemo, and Assoc. Prof. Dr. Rashidah Funke Olanrewaju,
aahraryk1992@gmail.com aisha@iium.edu.my
International Islamic University Malaysia, Gombak, Kuala Lumpur, Malaysia

Abstract
One of the objectives and aspirations of scientists and engineers ever since the development of computers has been to interact naturally with machines. Hence features of artificial intelligence (AI) like natural language processing and natural language generation were developed. The field of AI that is thought to be expanding the fastest is interactive conversational systems. Numerous businesses have created various Virtual Personal Assistants (VPAs) using these technologies, including Apple's Siri, Amazon's Alexa, and Google Assistant, among others. Even though many chatbots have been introduced through the years to diagnose or treat psychological disorders, we are yet to have a user-friendly chatbot available. A smart generative cognitive behavioral therapy with spoken dialogue systems support was then developed using a model Persona Perception (P2) bot with Generative Pre-trained Transformer-2 (GPT-2). The model was then implemented using modern technologies in VPAs like voice recognition, Natural Language Understanding (NLU), and text-to-speech. This system is a magnificent device to help with voice-based systems because it can have therapeutic discussions with the users utilizing text and vocal interactive user experience.
Keywords:
Machine Learning, Mental Health, Therapeutic Chatbot, Deep Learning Approaches, GPT-2.

1. Introduction

In January 2020, the World Health Organization (WHO) estimated that more than 264 million people worldwide experience depression [1]. Additionally, it stated that depression is the main contributor to disability and might result in suicide. Following the Coronavirus Disease 2019 (COVID-19) pandemic, WHO remarked in March 2022 that fewer than 2% of global health funds are allocated to mental health [2].

Currently, just 6% of 165,000 healthcare applications accessible in smartphone application stores are focused on mental health issues [3].

In Cognitive Behavioral Therapy (CBT), several cognitive and behavioral interactions are used as a well-known and scientifically validated treatment [4]. CBT's concept is based on the significance of false ideas and mindsets, improper information processing, and unhelpful behavior as the risk factors for depression [5]. Cognitive-behavioral approaches are therefore presented and practiced during treatment sessions with classwork to internalize the new behavior [6]. Because most of these sessions are conversation based, the CBT approach is suitable for this research. While keeping track of the patient's assignments and progress, the chatbot will record their dialogues. The assignments and progress reports help them modify their views and thinking through time. Figure 1 depicts the efficacy of CBT according to Kaur and Whalley, 2020 [6].

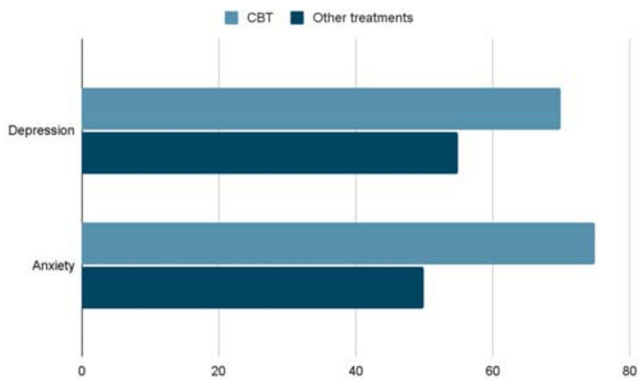

Figure 1: The effectiveness of CBT compared to other treatments.

Spoken dialog systems are recently finding their way into all intelligent devices and gadgets. It provided user-friendly, efficient, and human-like communication for the users. These technologies are implemented in education, government, business, and entertainment industries. Nevertheless, they are yet to prove their benefits, particularly in the mental health sector. In the world of Virtual Personal Assistants (VPAs), there are different methods. Every company

Manuscript received April 5, 2023
Manuscript revised April 20, 2023
https://doi.org/10.22937/IJCSNS.2023.23.4.3

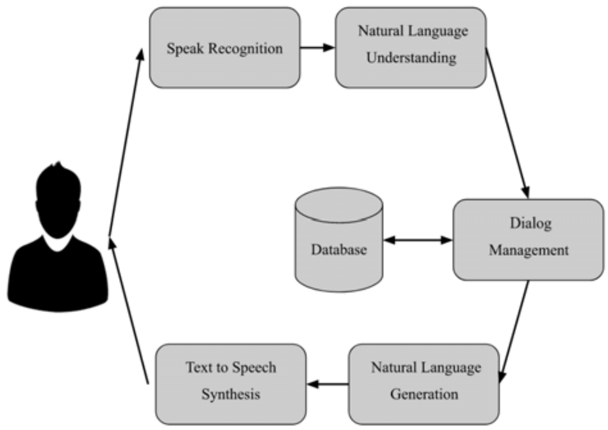

Figure 2: The Structure of Dialogue System

has its preferred method and implementation. For example, Google Assistant uses Deep Neural Networks (DNN) that generally focus on the main components of dialog systems [7]. On the other hand, Amazon is benefiting from Automatic Speech Recognition (ASR) methods and Natural Language Understanding (NLU), as mentioned on their website.

It is expected that during a typical conversation, each participant takes a turn to talk. The same is anticipated of a spoken dialog system, a natural and efficient interaction. There are six main components in each spoken dialog system, as presented in figure 2.

Generative conversational chatbots are designed to have a conversation with users. Generative chatbots' most attractive and unique feature is that they improve over time by obtaining past interactions. There are two approaches to designing conversational models, rule-based and machine-based [8]. Rule-based interactions are based on predefined rules, while machine-based is learning and improving over time by utilizing deep-learning techniques. Generative models fall in the machine-based category. While improving based on the question and past interactions makes them more innovative, it is also more prone to error. Training with larger datasets can help improve their accuracy [9].

2. Statement of significance

Everyone has a right to good mental health. Therefore, it is necessary to have access to psychological techniques to deal with these issues. The significant therapeutic choices are currently self-help, medical treatments, and psychotherapy.

Unfortunately, not everyone has access to or can afford to attend a psychotherapy session. And either of the chatbots available can have a vocal conversation with patients and mostly are quiz-like rather than letting them share their opinions.

This paper focuses on designing a system that allows patients to have meaningful human-like conversations with a cost-effective intelligent therapist chatbot to improve their depression using CBT.

3. Material and methods

The following sections present the details of achieving a mutual dialogue generation model. The dialogue generation selected for this study has five parts: transmitter model, supervised dialogue generation, fin-tuning the mode, reward shaping, and finally, receiver model. This section started with a detailed design flowchart. This flowchart is an important figure presenting this study's extensive yet detailed image. After that, the process of data acquisition is described. Gathering reliable yet accessible data was significant hence its direct impact on how the GIPE bot would respond to patients.

3.1. Detailed Design Flowchart

Figure 3 is a detailed flowchart of the GIPE design. This figure lays out a breakdown of each block and delivers a broad image of the whole research. This figure is divided into three sections: model training, dialogue generation, and user interface. The model training section presents an illustration and example of how raw data is processed, impeded, and finally used for training the model. Transmitter-Receiver machine learning algorithms are illustrated. With the help of the evaluation model, these algorithms are used for training the GIPE model. After the training phases, patients will interact with the model through the user interface. The dialogue generation section is the medium that helps process the user's input to prepare for interacting with the GIPE model and then return

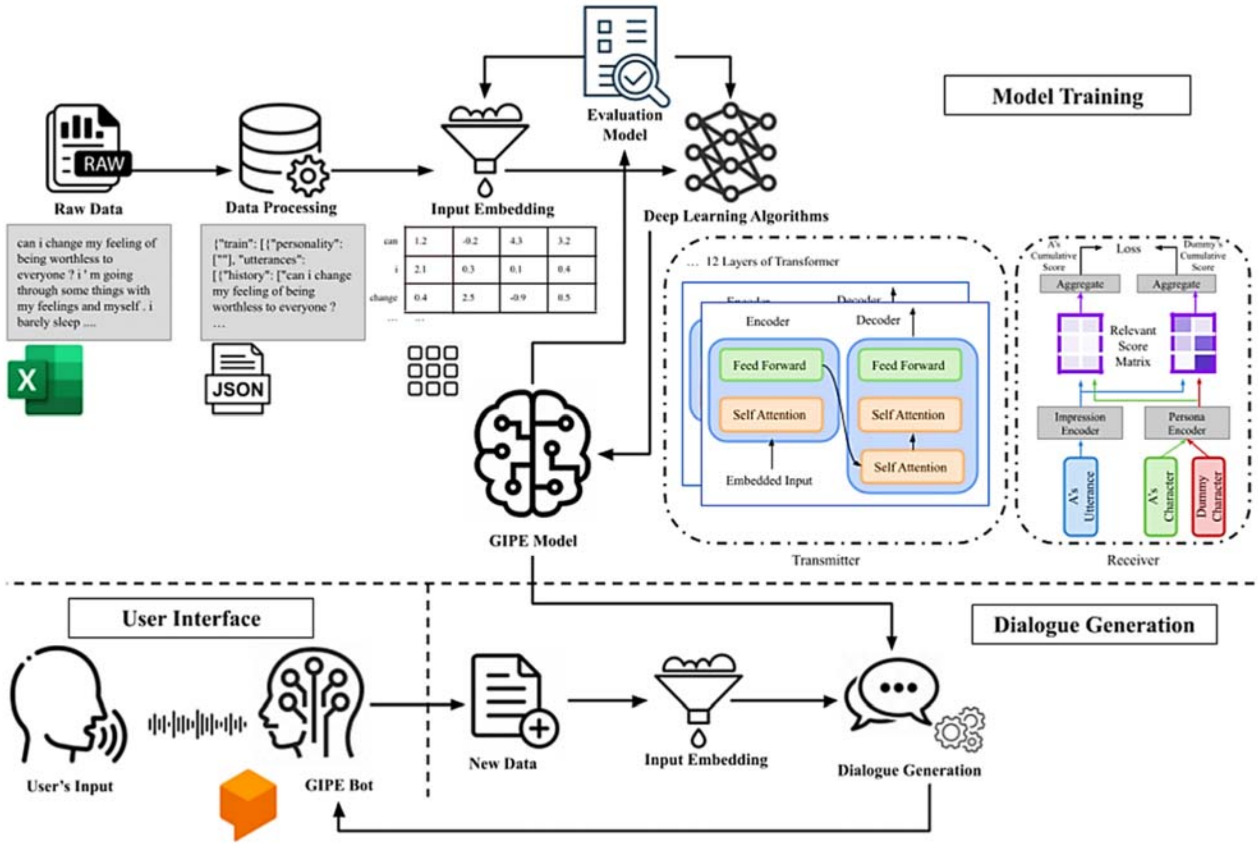

Figure 3. Proposed GIPE Bot's detailed architecture

the response. Users can interact with the GIPE bot via both voice and text input.

According to Li et al. [14], dialogue production is regarded as a sequence generation difficulty in this work. The pretraining transformer language model (GPT-2) published by Radford et al. [15] is utilized to initialize the transmitter. The two stages of training are supervised conversation production and fine-tuning of the self-play model. The supervised dialogue generation issue is optimized via maximum likelihood estimation (MLE). In the self-play model fine-tuning, reinforcement learning (RL) is used to urge the transmitter to develop a strategy that maximizes reward signals by simulating the discussions between two randomly assigned interlocutors. Language modeling and shared character perception are used in the reward function design.

3.2. Data Acquisition

It is particularly challenging to locate reliable statistics on mental health treatment. Moreover, the available data are a poor approximation of genuine contact between a patient and a therapist. Indeed, one can scrape Reddit, social news aggregation, and discussion websites and find intriguing therapeutic conversations between people. Still, it is almost impossible to verify the qualifications of self-claimed professionals. Moreover, most of the other datasets that are accessible are either costly or proprietary.

After some investigation, a collection of therapist replies to genuine patients was discovered on the Counsel Chat website in December 2020 [10], an open-source, relatively high-quality mental health query. Counsel Chat website is a form of an expert community. It serves as a platform for therapists to establish their credibility and connect deeply with prospective clients. On the website, therapists answer client inquiries, and visitors can "like" the most valuable answers. It is an intriguing concept that can produce some informative data.

The fact that verified therapists post these replies is an outstanding aspect of this information. Although they might not always be the best responses, they come from a subject matter expert. If Reddit data is utilized, anyone can be the one giving advice. Here, the people providing the guidance are licensed counsellors. It's crucial to remember that encounters with therapists in person can differ significantly from those that are made public online. Additionally, this is not a conversation between a therapist and a client. There is only one discussion involved at a time.

Bertagnolli et al [11] were able to get in touch with the website's founders and access some actual data from the website. There is data analysis on these data in section 4, and the dialogue generation model will be trained on these as well.

Since it is generally accepted that objective measurements and human assessment outcomes have poor correlations, human evaluation was further used in this study. Though human assessment on an extensive test dataset is expensive and difficult to compare with other models in the literature, it is nevertheless necessary for dialogue development. Semantic and topical similarities are measured using the average embedding metric. This metric indicates that a high score will be obtained if the semantic content of the model-generated answer and the regression coefficients response is similar. Following prior research by Serban et al [12] and using the embedding measurement to some extent indicate the response quality [13].

3.3. Data Conversation

The required signature for the JSON file containing the training data is shown in Figure 4. There are two primary keys in the more significant JSON object. "valid" and "train." A collection of characteristics and utterance pairings makes up the training data called the train. Except for the validation set, valid is the same. The speaker's characteristic is described in a series of phrases called the character.

The character field was left empty for this model to denote a lack of characteristics data. Next, a list of candidate replies may be found in the candidate's section. This list includes several less-than-ideal replies to the conversation's history, where the last statement is the actual answer. Finally, it is necessary to define history. The history is just a list of characters with a new conversation turn at each place.

This bot has two conversation turns, one from the individual asking the inquiry and one from the therapist in response. Thus, the question must be answered. This format is carried out because the model does not want to train a general therapist bot but wants to find appropriate singular answers to situations. The model can be used immediately if the data is correctly prepared.

```
{
 "train": [
   {
     "personality": [
       "sentence",
       "sentence"
     ],
     "utterances": [
       {
         "candidates": [
           "candidate 1",
           "candidate 2",
           "true response"
         ],
         "history": [
           "response 1",
           "response 2",
           "etc..."
         ]
       }
     ]
   }
 ],
 "valid": ...
}
```

Figure 4. JSON file template for training data

3.4. Developing an Interactive Voice Response System

After the consultation ML model is trained on the acquired data, and the model is evaluated, the final stage of this research is to develop a dialogue-spoken model application that can interact with patients via both text and voice.

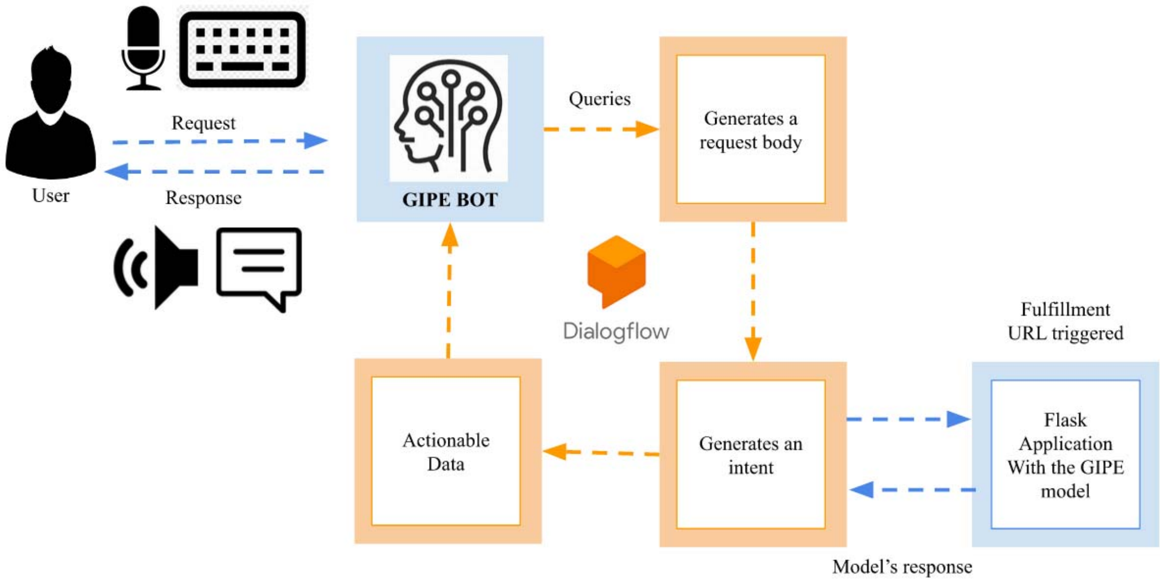

Figure 5. The proposed interactive voice response chatbot architecture

To develop the application, DialogFlow was employed. It uses Dialogflow. Dialog Flow is an NLU platform that uses conversational SLU techniques and CMU Sphinx (speech recognition systems). A conversational user interface can be designed and integrated into mobile apps, online applications, gadgets, chatbots, interactive voice response systems, and other applications. Dialogflow may examine a variety of user inputs, such as text or audio sources (like a voice recording). Additionally, it can respond to consumers through text or artificial speech.

The trained ML model is uploaded to a public HTTPS URL to achieve this. This application flask allows direct interaction with the model. Then after creating a project in Dialogflow, intents are linked to the Flask application. Figure 5 presents this system's architect for a better visual understanding.

4. Results

Firstly, several charts are introduced to visualize different aspects and frequencies of the data. This section is then followed by some raw examples of users' conversations with the consultation model, and the mobile application prototype.

Table 1. Data Analysis Columns

No	Column header	Description
1	questionID	A particular question identifier that is different for every question.
2	questionTitle	The question's title on the counsel chat page.
3	questionText	The bulk of the client's inquiry to the counsellor.
4	questionLink	The most recent URL for the query.
5	topic	The heading of the question was listed under it.
6	therapistInfo	A description of each therapist, typically including their name and areas of expertise.
7	therapistURL	A connection to the counsellor's profile on counsel chat answer.
8	answerText	The therapist's answer to the query.
9	upvotes	The number of upvotes for the answer.
10	views	How many times has the question been viewed.
11	split	Divided data into training, validation, and testing groups.

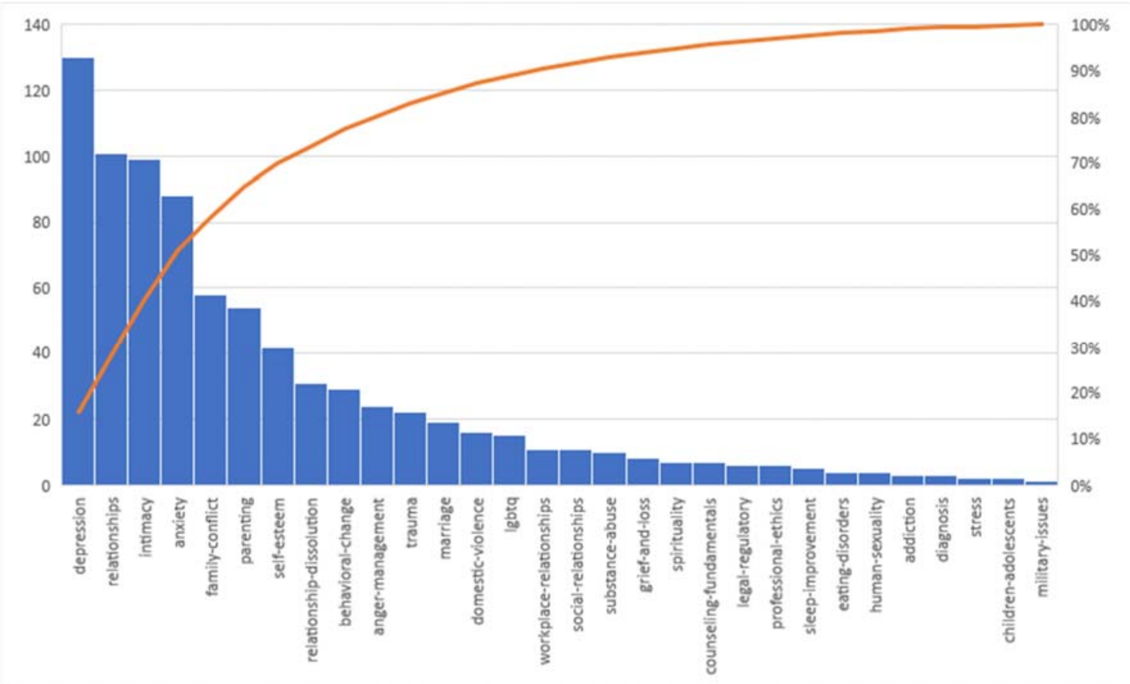

Figure 6. Number of questions per topics

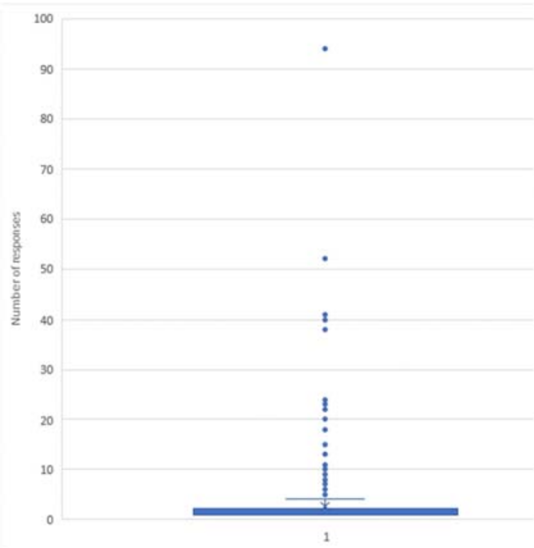

Figure 7. Number of responses per question

4.1. Dataset Analysis

This data analysis is based on data provided by the Counsel Chat website, 25 March 2020 [10], The dataset is shown as a CSV file with eleven columns. Table 1 presents the columns headers and their descriptions.

There are 818 unique questions, and overall, 2,129 responses to them. The questions are on 30 topics, including depression, anxiety, and self-esteem, figure 6. As the figure presents, depression covers the most of questions with about 130 questions, following by relationships and intimacy with each approximately 100 questions. At fourth place there is anxiety with 90 questions. Hence the topics of depression and anxiety are in the top four with over 220 questions, this dataset is valuable for the purpose of this study.

Figure 7 shows the number of responses per question. About 75% of questions have two or fewer responses by there are questions that are highly engaged therapists. The most responded to the question is, “Do I have too many issues for counseling?”.

Most questions are short, and therapists seem to provide much more extended answers. For example, the average length of a question was about 64 words, and the average of a therapist’s answers was about 170 words, Figures 8 and 9. Having longer answers assists in training the GIPE bot to produce longer responses

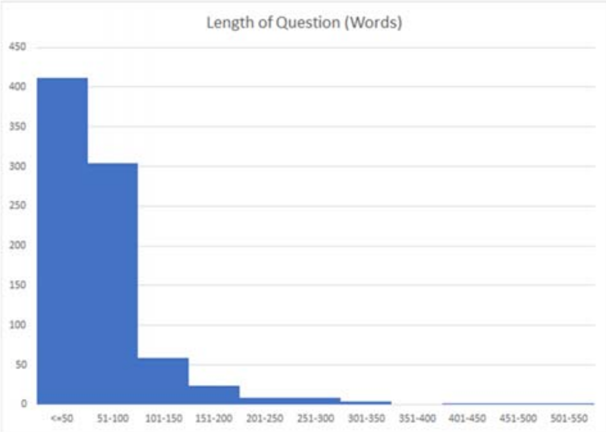

Figure 8: Distribution of length of questions

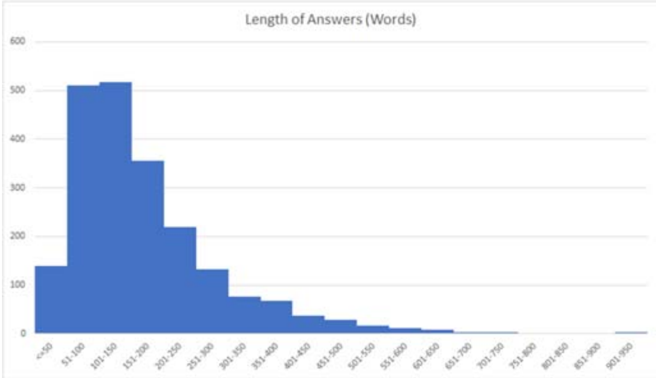

Figure 9: Distribution of length of responses

as well which would result in a more descriptive and informative interaction with the patients.

Figure 10 represents the number of answers per topic. Depression, relationships, intimacy, and anxieties are the top questioned topics in order. Interestingly consulting fundamentals found its place as the top 3 most answered topics. Answers related to depression were about 340 in the first place. Second place is answers for anxiety related topic with almost 250 answers. Together these two topics have approximately 590 answers. Hence GIPE both generate responses, having a higher number of responses would result in more diverse and comprehensive answers. Hence, this dataset is perfect for GIPE bot training.

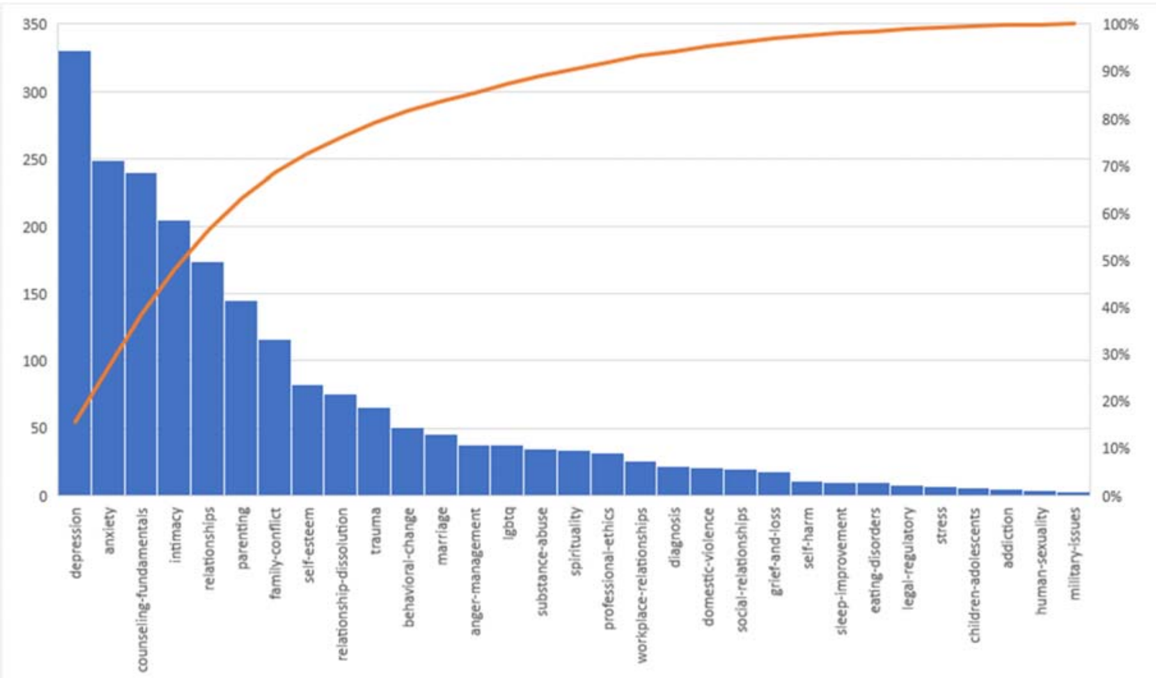

Figure 10. Responses counts for each topic.

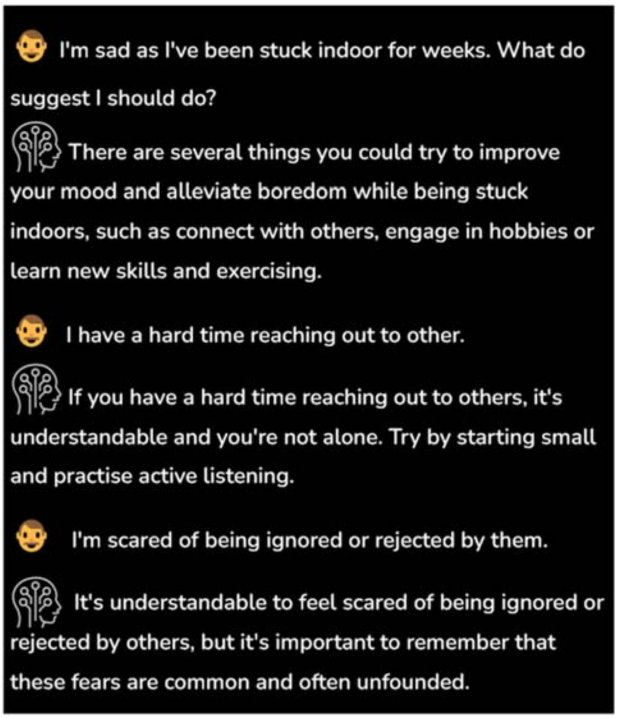

Figure 11. Example of the GIPE chatbot conversation with patient

4.2. Consultation Model

After training the dialogue generation model with the dataset acquired, the consultation chatbot was tested with some simulated conversations. Figure 11 shows the consultation chatbot's conversation with the user. As observed, the chatbot is empathetic and provides relevant information and techniques to help the user. As presented in the example, the chatbot shows awareness; hence the responses sound human-like and relevant. Also, it can remember and reuse the information provided earlier. This promising feature illustrates the generative ability of the GIPE bot. As expected of an expert chatbot, the responses are informative and emphatic.

While these results seem highly promising, the true potential of the GIPE chatbot can only be tested in a clinical environment and with actual patients. In the presence of psychologists and consultants, patients' current mental health should be established. Then they will have regular interaction with the GIPE chatbot. After a fixed set of time, their mental health should be examined. This approach is the only way that the true impact of the GIPE bot can be fully understood and measured.

4.3. Mobile application

The final stage of this study was developing a user-friendly mobile application. This application can keep the flow of natural conversations with patients diagnosed with depression and recommend techniques to overcome it. We are also hoping to reach youth around the globe without access to mental healthcare. The mobile application's prototype is presented in Figure 12. As shown, users can both type their messages to the GIPE bot or send a voice note which then will be converted to text. The design is simple and direct to not cause any confusion for the users as well as to bring the learning curve to the lowest possible. A record of conversions is kept on the user's device for future reference.

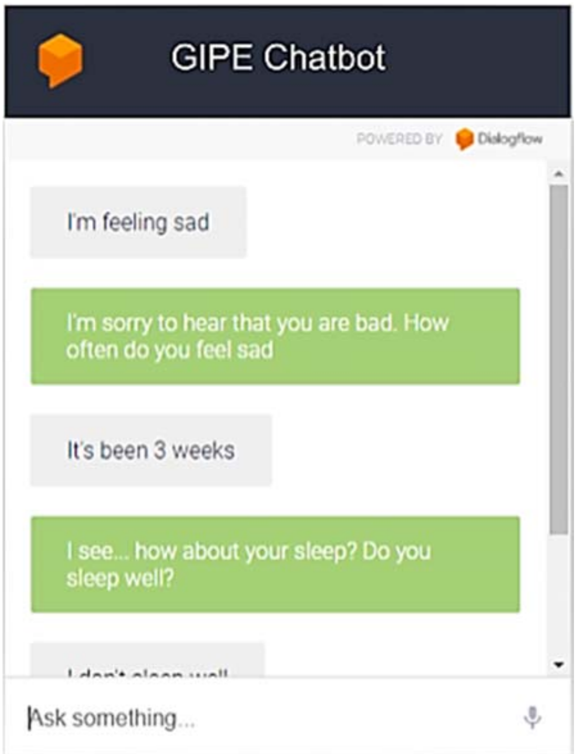

Figure 12. The mobile application prototype.

5. Discussion

This study provided data analysis based on information provided by the Counsel Chat website, 25 March 20220, to Bertagnolli et al., 2020. There are 818 unique questions, and overall, 2,129 responses to them. About 75% of questions have two or fewer responses by there are questions that are highly engaged therapists. The most responded question is, "Do I have too many issues for counselling?". Then the data conversion was achieved by the two primary keys in the more significant JSON object, "valid" and "train." A collection of characteristics and utterance pairings makes up the training data called the train. This collection is just a list of characters with a new conversation turn at each place. The character field was left empty for this model to denote a lack of characteristics data. After that, the dialogue-generation machine-learning model was trained on the processed dataset.

6. Conclusions

The paper aimed to design a model that allows users access to an affordable cognitive therapist chatbot that provides expert advice on mental health. The following presents the achievements of this study concerning the objectives selected.

A model was selected and was connected to a user-friendly mobile application with a simple yet effective user interface. The mobile application design allows users of different technological backgrounds to use it with only a few minutes of training.

Finally, the model was connected to the Dialogflow system. This stage provides the interactive voice feature with a user-friendly user interface. Users have the option of interacting with the GIPE bot via both voice and text messages. The mobile application's design is simple and to the point, eliminating any confusion. Any user with limited knowledge of technology can use it within a few minutes of tutoring.

A conclusion chatbot developed during this study with high-quality, real-life consultation responses from psychology professionals resulted in a chatbot with traces of empathetic and human-like repose. Furthermore, the replies are informative, which is an excellent promise in case of support for testing in a clinical trial setting. Finally, a simple and user-friendly mobile application is designed for people of different technological backgrounds.

Its impact can only be tested after a clinical trial for a psychology-related work like this. In the future, involving psychologists to evaluate the model's response similar to the human evaluation technique but considering the responses of the model as psychologically correct and informative can help improve the model tremendously. Also, involving patients can provide the most critical feedback for improving the design.

In training, an expert dialogue generation model such as GIPE, a more significant and diverse number of datasets, can help improve the chatbot's responses to be more informative and expert sounding. Also, data from professionals with different techniques can diversify the chatbot's replies and reactions.

Another implementation to be considered is connecting this chatbot to a physical device to smaller devices such as smartwatches or voice-activating devices that can help patients reach out anytime needed without carrying their phones.

References

[1] Depression [Internet]. World Health Organization. World Health Organization; [cited 2020Dec19]. Available from: https://www.who.int/news-room/fact-sheets/detail/depression

[2] Covid-19 pandemic triggers 25% increase in prevalence of anxiety and depression worldwide [Internet]. World Health Organization. World Health Organization; [cited 2022Apr6]. Available from: https://www.who.int/news/item/02-03-2022-covid-19-pandemic-triggers-25-increase-in-prevalence-of-anxiety-and-depression-worldwide

[3] Carlo AD, Hosseini Ghomi R, Renn BN, Areán PA. By the numbers: ratings and utilization of behavioral health mobile applications. NPJ digital medicine. 2019 Jun 17;2(1):1-8. doi: 10.1136/ebmental-2018-300069

[4] Hofmann SG, Asmundson GJ, Beck AT. The science of cognitive therapy. Behavior therapy. 2013 Jun 1;44(2):199-212. doi: 10.1016/j.beth.2009.01.007

[5] Butler AC, Chapman JE, Forman EM, Beck AT. The empirical status of cognitive-behavioral therapy: a review of meta-analyses. Clinical psychology review. 2006 Jan 1;26(1):17-31. doi: 10.1016/j.cpr.2005.07.003

[6] Jurinec N, Schienle A. Utilizing placebos to leverage effects of cognitive-behavioral therapy in patients with depression. Journal of affective disorders. 2020 Dec 1;277:779-84. doi: 10.1016/j.jad.2020.08.087

[7] Këpuska V, Bohouta G. Improving wake-up-word and general speech recognition systems. In2017 IEEE 15th Intl Conf on Dependable, Autonomic and Secure Computing, 15th Intl Conf on Pervasive Intelligence and Computing, 3rd Intl Conf on Big Data Intelligence and Computing and Cyber Science and Technology Congress (DASC/PiCom/DataCom/CyberSciTech) 2017 Nov 6 (pp. 318-321). IEEE. doi: 10.1109/DASC-PICom-DataCom-CyberSciTec.2017.67

[8] Khalaf AA, Hashim AH, Olowolayemo A, Funke R. Artificial Intelligent Applications for Mental Health Support: A Review paper. Engineering Professional Ethics and Education 2021 (ICEPEE'21). 2021 Jun 22:22.

[9] Varghese E, Pillai MR. A standalone generative conversational interface using deep learning. In2018 Second International Conference on Inventive Communication and Computational Technologies (ICICCT) 2018 Apr 20 (pp. 1915-1920). IEEE. doi: 10.1109/ICICCT.2018.8473211

[10] CounselChat. Mental health answers from counselors [Internet]. CounselChat. [cited 2020Dec]. Available from: https://counselchat.com/

[11] Bertagnolli N, Lord G, Ström E, Lee P. Counsel chat: Bootstrapping high-quality therapy data [Internet]. 2020 [cited 2020Dec]. Available from: https://towardsdatascience.com/counsel-chat-bootstrapping-high-quality-therapy-data-971b419f33da

[12] Serban I, Sordoni A, Lowe R, Charlin L, Pineau J, Courville A, Bengio Y. A hierarchical latent variable encoder-decoder model for generating dialogues. InProceedings of the AAAI Conference on Artificial Intelligence 2017 Feb 12 (Vol. 31, No. 1).
doi: 10.1609/aaai.v31i1.10983

[13] Serban I, Klinger T, Tesauro G, Talamadupula K, Zhou B, Bengio Y, Courville A. Multiresolution recurrent neural networks: An application to dialogue response generation. InProceedings of the AAAI Conference on Artificial Intelligence 2017 Feb 12 (Vol. 31, No. 1).
doi: 10.1609/aaai.v31i1.10984

[14] Li, J., Monroe, W., Ritter, A., Galley, M., Gao, J., & Jurafsky, D. (2016-a). Deep reinforcement learning for dialogue generation. arXiv preprint arXiv:1606.01541. doi: 10.48550/arXiv.1606.01541

[15] Radford, A., Wu, J., Child, R., Luan, D., Amodei, D., & Sutskever, I. (2019). Language models are unsupervised multitask learners. OpenAI blog, 1(8), 9.