

Virtual Assistant Chatbot for Mental Healthcare

Priom Deb
Computer Science and Engineering
School of Data and Sciences
BRAC University
Dhaka, Bangladesh
priom.deb@g.bracu.ac.bd

Asibur Rahman Bhuiyan
Computer Science and Engineering
School of Data and Sciences
BRAC University
Dhaka, Bangladesh
asibur.rahman.bhuiyan@g.bracu.ac.bd

Md Sabbir Hossain
Computer Science and Engineering
School of Data and Sciences
BRAC University
Dhaka, Bangladesh
md.sabbir.hossain1@g.bracu.ac.bd

Md. Farhadul Islam
Computer Science and Engineering
School of Data and Sciences
BRAC University
Dhaka, Bangladesh
md.farhadul.islam@g.bracu.ac.bd

Annajiat Alim Rasel
Computer Science and Engineering
School of Data and Sciences
BRAC University
Dhaka, Bangladesh
annajiat@gmail.com

Abstract—Early researchers have tried using chatbots to provide users with psychological treatment. Through dialogues, ongoing user monitoring, and ethical consideration during the intervention, the program does not take the user’s psychiatric health into account. We suggest that consumers who require mental health care are more satisfied when emotion recognition is more accurate and ongoing. We tried to make a chatbot for psychiatric counseling that uses high-level NLU. The approaches allow for delicate, ongoing study of emotional changes. Additionally, a relevant response to clinical psychiatric counseling is offered by the case-based counseling response model that incorporates an ethical judgment model.

I. INTRODUCTION

The ability of computers to recognize human emotions has long been a subject of study. Numerous studies have recently improved human emotion recognition using artificial intelligence (AI) techniques. The research uses a large amount of emotion-labeled data to train a variety of deep learning-based emotion classification models, including convolution neural network [6], recurrent neural network, and attention network [7]. The research presents important findings about emotion recognition. They are quite good at identifying human emotions. In this paper, we describe a chatbot application for counseling that uses the chat assistant platform and the aforementioned emotion identification techniques to offer conversational services for mental health care. Due to improvements in public perception of mental care, many patients are now receiving psychotherapy. Despite the fact that many people nowadays suffer from mental diseases, there is still a low incidence of actual diagnosis and treatment. Since it might be difficult to recognize one’s own mental illness, this is true generally. Because there aren’t enough professionals to meet demand, consulting with experts like psychiatrists, mental clinical counselors, and counseling psychologists is pricey. Anyone can readily obtain and receive effective treatments for mental counseling services by using the conversational service. The conversational service can offer

individuals one-on-one, individualized counseling. It’s critical to end the isolation of people suffering from mental disorders including sadness and lethargy [10]. A one-on-one discussion can successfully end the isolation. When a patient requires urgent assistance, personal conversations can also be effective [11]. When an accidental mental disorder, such panic or a suicide impulse, occurs, the service alerts the user of their risky status. Suicidal people require an immediate and prompt response. It is possible to properly notify those who are close to the user and take actions to help prevent suicidal impulse. The most crucial aspect of psychiatric counseling is linguistic contact. The encounter is a standard instance of insight-oriented psychotherapy, a psychoanalytic treatment method popularized by Freud. Through verbal contact, patients can better comprehend their own issues in order to discover a solution. To make it simple to use the counseling service, we use the chat assistant platform. A chatbot that can be used on different messaging platforms as a conduit between users and conversational services is included in the chat assistant. The contents of the consultation can be typed by people who want psychiatric counseling. Here, we apply a variety of NLP techniques to assess the contents of consultations. The service can also accept a variety of extra data from wearable sensors, picture, video, and audio sources, among others. The chatbot discerns the user’s current emotions and reports its findings to the conversational service. The service displayed outcomes logs as well as emotional highs and lows. Through such continual conversations, we continuously track and diagnose the psychological issues that consumers are experiencing. After that, the service uses natural language generation (NLG) techniques to respond correctly to the user’s fresh inputs. The service notices when a user’s fragile emotional state changes and responds immediately to handle crucial situations. In the section that follows, we examine related approaches to creating conversational services. The chatbot for psychiatric treatment is used using some of the findings from the analysis

of dialogues and natural language understanding.

II. BACKGROUND

A. Emotion Understanding from Several Data-sets

To enhance the communication between people and machines, emotion recognition is a crucial area of research. The task of acquisition is made more challenging by emotional complexity. A user's emotion can be determined using a range of data. Text [8], images [6], video [6], audio [7], and so on all include emotional content. Reference [6] provided a committee-based system. Deep CNN machines and their application in robust FER are demonstrated using SFEW 2.0 competition statistics. In Reference [7], a method was used to process the audio, video and video data. The emphasis of Reference [9] was on automatically assessing a movie and the speaker's attitude toward internet movie reviews. They thought about fetching media features for sealing the important information. There have been numerous studies on emotion analysis using several modalities. Our goal is to create an emotion-filled machine. To begin realizing the goal, a device or system that can think like a human being, feel warmth in the soul, judge events, prioritize options, and use many more emotional terms is required.

B. AI-Powered Virtual Assistant in Mental Health Care

The virtual assistant is a computer program that can interact with the human through dialogues and conversations. The conversation can be using texts, voice commands, or a mixture of both mediums. The virtual assistant should be able to interpret human emotions and responses in a psychiatric manner. To achieve this ability in a virtual assistant the program should understand the natural language and respond using the natural language generation method. Therefore, both Natural Language Processing and Natural Language Generation is the most crucial part of a virtual assistant in mental health care. There are several research on that work on such virtual assistants (also known as chatbots). To give an overview seven different databases were used in the research [3]. In that paper, they describe a list of fields where the chatbots were used for. Among all the lists they found that the most focus on chatbots is in case of depression. The current development in mental health care chatbots is still immature. Different types of chatbots are being developed and used nowadays. Therefore, AI-powered virtual assistants are still the most desired study among NLP researchers around the globe especially when it comes to mental health. To understand the user's emotions it is necessary to mimic the chatbots that will act like the user's friends and family or even a psychiatrist. The chatbot should not reply based on some pre-ready responses. Rather, it should be able to generate responses with proper reasoning. When the user interacts with the chatbot, the chatbot should be able to interpret the current state of mind of the user. Then, after understanding the state of mind, the chatbot should respond with proper reasoning and other aspects to give mental relaxation to the user. Moreover, while interacting with chatbots, facial expressions can be taken into consideration to

interpret the current emotional state of a user more precisely. Here deep learning model can be used to understand facial expressions. Finally, learning should be a continuous process for the chatbot to understand the user.

III. NATURAL LANGUAGE UNDERSTANDING

A. Word Sense from Sentences

The virtual assistant needs to respond to the user adequately. For this reason, the virtual assistant should be enriched linguistically and in vocabulary. The virtual assistant should be able to continue the conversation with the user to understand the emotional state of the user. It should not only respond based on psychiatric sentence generation. But sometimes it should respond in a way so that the user understands that the virtual assistant is actually listening and understanding him/her. To accomplish such ability, paraphrasing can be an important method.

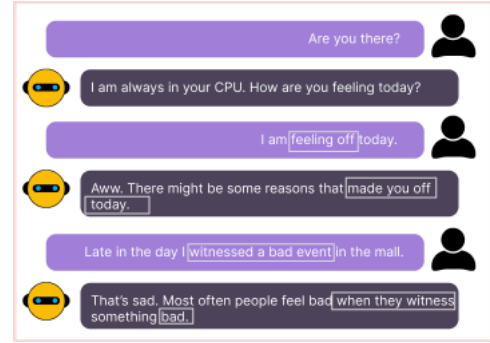


Fig. 1. An example of paraphrasing.

In figure 1, we show a simple example of paraphrasing where the bot continues the conversation with the user. To paraphrase a sentence the chatbot must be trained enough to generate similar sensed sentences. There are several methods to achieve this paraphrasing functionality. But based on a research, the word embedding model can be a useful method to achieve paraphrasing functionality in virtual assistants [2]. According to their research, first, we need to gather a series of corpora of desired languages. The word embedding model helps to incorporate semantic knowledge within the virtual assistant. With the use of this model, we can train the machine to understand the similarity and lexical meanings of different words. We use word vectors to find the similarity to replace a word with another appropriate word. The correctness of the similarities depends on the size and collections of the corpus. To get precise results, we need a large scaled corpus. There are several ways to find the similarities between words. Using cosine to compute the word similarities is one of the most efficient ways in word embeddings. Cosine values can be used to find word similarities [4]. When cosine calculation of the word vectors is -1 [4], then the word vectors are in the opposite direction. On the other hand, when cosine calculation of the word vectors is +1, then the word vectors are in the same semantic direction [4]. Lastly, if the cosine

value is 0, the vectors are orthogonal. But raw frequencies are positive and that is why the cosine values range from 0 to 1. Therefore, when the cosine value is higher there is a similarity between the words. We can replace one word with another high-similarity word to paraphrase a sentence. This word embedding model will help to generate sentences to continue the conversation between the virtual assistant and the user.

B. Sentence Context and Sentential Behaviour Classification

When user enters his/her dialogues, it should be scanned in order for the chatbot service to appropriately respond to the user's emotions. By using a model of language similarity analysis based on GRU [12], we examine the domains with the highest relevance to the sentences provided. One can calculate the distance between domains using the vector. It also analyzes the utterance aim to determine if it is a broad discussion or a one liner to know the right answer.

C. Analyzing Spatial-Temporal Context

To understand the user after a series of conversations, the chatbot should be able to understand the context of the conversation. The chatbot should have a memory to stack some keys from the previous dialogues. It will help to backtrack and connect some words in future responses. After analyzing the spatial-temporal context, the chatbot can reply with some meaningful responses. Spatial-temporal context analysis is one of the most crucial functionalities for mental health care virtual assistants.

IV. ANALYZING EMOTIONAL INPUTS

The virtual assistant needs to understand the user behavior. For this reason, it is required to analyze the inputs that is provided by the users. The input data can be in different media such as audio (speech recognition), video/image (facial recognition). Using those data from the users the virtual assistant can analyze the current state of mind of the user. This analyzing mechanism is an important task of a virtual assistant for mental treatment.

A. Modeling Emotional Expressions

To understand human emotions, we need to represent all human emotions in a proper data model. Plutchik defined 8 major emotions: anger, fear, sadness, disgust, surprise, anticipation, trust, joy [5]. Also, some other emotions are correlated with the primary eight emotions. The chatbot should understand and be able to recognize all eight types of primary emotions. Besides, the chatbot needs to recognize and learn most of the other emotions that are correlated with the primary relations. After recognizing the emotions of the user, the chatbot can continue for sentence generation.

B. Training Data for Emotion Recognition

As we are targeting our virtual assistant to be able to interpret human emotions in different situations, we need a large amount of data that contains emotional information. The data could be collected from different media like TV shows,

speech, radio, paper, etc. The larger data we collect, the larger corpora can be built to precise the emotion recognition by the virtual assistant.

C. Recognizing and interpreting emotions

The most of the facial expression mechanism is rely on a collection of image data. Visible images, however, are easily impacted by changes in lighting, and posed expressions look and act differently from those of real people. In this section, a method is used to assess and infer a user's emotions by combining their mutual utterances in text, voice, and body language and details about images [6]. The reference [1] uses some useful models to recognize emotions from face recognition.

D. Continual emotional observation

The detection of experiencing mood swings is very important in mental therapy to enhance the performance of chatbot. We apply the patient mood changing technique using lifelogging. We apply the user emotion tracking technique using lifelogging. Keeping up with the System should be designed for data sets collection. This makes it possible to improve emotional health using physiological signal sensors and emotional data models for identification.

V. SENTENCE GENERATION FOR PSYCHIATRIC COUNSELLING

A. Emotional Expression and Communications

While expressing emotions and response to the users the virtual assistant need to consider a list of information. First of all, the age of the user should be considered while generating responses. After that, gender distinction is another important classification to generate mental health advice. To implement a more accurate response there is plenty of information could be collected from the user. But for the simplistic approach of implementation only age and gender could be used. Now the challenge comes to generate an appropriate response based on the user's current emotional state of mind.

B. Psychiatric Intervention

Virtual assistants for mental health care need to understand human emotion. Because based on their understanding skill the response is dependent. Then, the virtual assistant can give a piece of advice or treatment in a psychiatric manner. It can respond with some advice to give treatment for major mental health problems. These problems can be depression, suicidal attempts, fear, stress, etc. With the proper treatment or advice from the virtual assistant, the number of these mental health problems can be reduced. Such virtual assistants can lessen the mental stress of the users.

C. Generating responses using a pointer network model

One of the fundamental components of the Chatbot system is response generating. Commercial chatbot systems have traditionally relied heavily on methods based on rules and templates. The chatbot does not respond to many scenarios

when we employ these strategies. You can get a completely different answer if you phrase your query somewhat differently. We provide a model that generates answer sentences using a machine learning approach to address this issue. Using the Point Network Model, the answer generation model can learn critical sentence-level keywords [10]. Using feature maps and a sentence corpus, the sentence's attention is learned. The model was quite successful in finding words to answer questions in various Q&A applications, and it provided a significant response in psychiatric counseling.

D. Moral Action and Conduct

The standards for evaluating human moral values should be studied collectively. The user's quality of life can be directly impacted by psychiatric counseling [13]. Therefore, the best course of action should be followed ethically when responding to or offering intervention. The human ethics code model will be applied to unethical circumstances if the technology is cutting edge.

VI. CONCLUSION

We introduce a chatbot for mental healthcare in this paper. In conversation, the chatbot helps mental disorder patients by giving suggestions. The chatbot interacts with a patient and recommends best solutions.

The service uses a variety of emotional intelligence techniques, such as multilanguage emotion recognition and facial expression, artificial intelligence such and long-term monitoring etc. which is necessary to interpret the current state of the mind of the users. The mechanisms helps to monitor the emotional fluctuations of users. Finally, therapy is more successful, which raises user satisfaction among those in need of mental health care.

REFERENCES

- [1] Wang, S., Liu, Z., Lv, S., Lv, Y., Wu, G., Peng, P., Chen, F., & Wang, X. (2010). A natural visible and infrared facial expression database for expression recognition and emotion inference. *IEEE Transactions on Multimedia*, 12(7), 682–691. <https://doi.org/10.1109/tmm.2010.2060716>
- [2] Oh, K.-J., Lee, D., Ko, B., & Choi, H.-J. (2017). A chatbot for psychiatric counseling in mental healthcare service based on emotional dialogue analysis and sentence generation. 2017 18th IEEE International Conference on Mobile Data Management (MDM). <https://doi.org/10.1109/mdm.2017.64>
- [3] Abd-alrazaq, A. A., Alajlani, M., Alalwan, A. A., Bewick, B. M., Gardner, P., & Househ, M. (2019). An overview of the features of Chatbots in Mental Health: A scoping review. *International Journal of Medical Informatics*, 132, 103978. <https://doi.org/10.1016/j.ijmedinf.2019.103978>
- [4] Jurafsky, D., & Martin, J. H. (2021). *Speech and language processing: An introduction to natural language processing, computational linguistics, and speech recognition*. Pearson Prentice Hall.
- [5] R. Plutchik, (2002). *Emotions and Life: Perspectives from Psychology, Biology, and Evolution*. Washington, DC: American Psychological Association.
- [6] B. K. Kim, J. Roh, S. Y. Dong, and S. Y. Lee, "Hierarchical committee of deep convolutional neural networks for robust facial expression recognition," *Journal on Multimodal User Interfaces*, pp. 1-17, 2016.
- [7] L. Chao, J. Tao, M. Yang, Y. Li, and Z. Wen, "Audio Visual Emotion Recognition with Temporal Alignment and Perception Attention," *arXiv preprint arXiv:1603.08321*, 2016.
- [8] H. Lee, Y. S. Choi, S. Lee, and I. P. Park, "Towards unobtrusive emotion recognition for affective social communication," In *proc. of 2012 IEEE Consumer Communications and Networking Conference*, pp. 260-264, 2012.
- [9] M. Wollmer, F. Weninger, T. Knaup, B. Schuller, C. Sun, K. Sagae, and L. P. Morency, "Youtube movie reviews: Sentiment analysis in an audio-visual context," *IEEE Intelligent Systems* No. 28 Vol. 3, pp. 46-53, 2013
- [10] T. Mikolov, K. Chen, G. Corrado, and J. Dean, "Efficient estimation of word representations in vector space," *arXiv preprint arXiv:1301.3781*, 2013.
- [11] D. Elmasri, and A. Maeder, "A Conversational Agent for an Online Mental Health Intervention," In *proc. of International Conference on Brain and Health Informatics*, pp. 243-251, 2016.
- [12] K. Cho, B. Van Merriënboer, C. Gulcehre, D. Bahdanau, F. Bougares, H. Schwenk, and Y. Bengio, "Learning phrase representations using RNN encoder-decoder for statistical machine translation," *arXiv preprint arXiv:1406.1078*, 2014.
- [13] L. W. Roberts, J. Battaglia, and R. S. Epstein, "Frontier ethics: Mental health care needs and ethical dilemmas in rural communities," *Psychiatric Services*, 1999.