PRERONA: Mental Health Bengali Chatbot for Digital Counselling

Chapter · February 2021

DOI: 10.1007/978-3-030-68449-5_28

6 authors, including:

Jia Uddin
Woosong University
130 PUBLICATIONS **489** CITATIONS

SEE PROFILE

PRERONA: Mental Health Bengali Chatbot For Digital Counselling

Asma Ul Hussna[1], Azmiri Newaz Khan Laz[1], Md. Shammyo Sikder[1], Jia Uddin[2], Hasan Tinmaz[2], and A.M.Esfar-E-Alam[1]

[1] BRAC University, Dhaka, Bangladesh
{asma.ul.hussna,azmiri.newaz.khan,md.shammyo.sikder}@g.bracu.ac.bd,
esfar.alam@bracu.ac.bd
[2] Technology Studies, Endicott College, Woosong University, Daejeon, Korea
jia.uddin@wsu.ac.kr,
htinmaz@endicott.ac.kr

Abstract. These three words "Human", "Machine" and "Interaction" refer to the modern era of technology where researchers emphasized the interaction between computer machines and humans. We humans cannot confine ourselves to relying solely on human-based healthcare systems. The current era of digitization brings advances and additional opportunities for medical care, especially in general and clinical psychology. Existing chatbots such as Woebot, Wysa, Moodkit, etc. do not allow users to express themselves. However, our chatbot PRERONA allows the user space to talk about whatever they want along with it is intelligent enough to ask questions as well as answer questions. In our research work, we have presented a model named 'PRERONA: Mental Health Bengali Chatbot for Digital Counselling'. PRERONA is designed to help depression struggling people especially for Bengali users who are lack attention and have none to talk about their problem. The main objective of this chatbot PRERONA is to give instant answers to questions and queries as well as to provide proper mental health care to Bengali's in times of depression. Although our bot PRERONA is made only for Bengali users, we have it in multiple languages such as Bengali, English and Korean. In addition to that, we are using adaptive learning process. If the bot can not answer any questions, that will be recorded as well as added to that database later on for more accuracy. Natural Language Processing (NLP) is used here to successfully implement these languages. The complete structure of this chatbot and how it works are observed accurately throughout this research.

Keywords: NLP, Interaction, Psychology, Adaptive, Depression, Mental Health Care.

1 Introduction

Machine learning has played a crucial part in the research of computational learning theory. It becomes an integral part of artificial intelligence to construct

algorithms when a computer program simulates human conversation through voice commands or text chats or both, we call that a chatbot [11] [7]. This automated program interacts with users like a human. Chatbot can play significant role extensions to health care services. If we specify, the need for chatbots in the psychological or mental health sector is rising along with the alarming rate of increasing number of depression and anxiety patients [4]. According to the world health organization (WHO) research, over 300 million people are struggling from depression worldwide which is equivalent to 4.4% of the world's total population. It is more common in females (5.1%) than males (3.6%) [12]. In Bangladesh, according to the first national survey on mental health (2003-2005), about 5 million people are suffering from depression including 4.6% adult and 1% children [5]. As per a recent WHO study, it is estimated that in Bangladesh, 4.1% of the population suffer from depression, and 4.4% from anxiety disorder [6]. But this is even more shocking that in spite of the existing treatments for mental disorders, 76%-86% people of low to a middle-income country do not seek medical attention for their disorders either out of negligence or hesitation due to social stigma. In this situation, our mental health chatbot PRERONA can play a vital role. It can ask real time questions to determine and fix the mental condition of the users, PRERONA lets its users share their feelings and also responses accordingly. Additionally, it provides the users with a lot of contents relating anti-depression as well as anti-anxiety. Moreover, PRERONA is a natural language processing (NLP) based tri-language chatbot, it is able to understand all languages based on the given database. Initially, we have set Bengali, English and Korean Corpus to our database which makes our bot more interesting. The main target behind PRERONA is to develop an online-based mental health system which can be considered an alternative to the typical existing systems. Here users will get 24/7 support and can express themselves without any hesitation along with proper guidance to get over their problems under expert assistance. This paper presents a detailed analysis of the proposed model of PRERONA, experimental setups and result analysis.

2 Literature Review and Related Work

Eliza was created in an AI laboratory in MIT between 1964-66 is the oldest and well known chatbots in the history of chatbots. Eliza was made in that way so that she could ask open questions with which she also answered, which simulated her role as a Rogerian psychotherapist [1] [10]. Now a days we can find a number of chatbots in practice such as Alexa, SIRI has brought revolutionary change in this field. Advanced technology has allowed for using chatbots in the creation of software-based therapy. Now the chatbot will see you instead of a therapist. Early work on mental health chatbot named Wysa and Woebot claims to be AI powered, follows cognitive behavioral therapy (CBT) [13]. If we consider works on Bengali chatbot, there are Golpo and doly which are NLP based and successfully able to chat with users in Bengali [9] [3]. Here, PRERONA has the features of both the types discussed above. It is a Natural language process (NLP) based

mental health chatbot which can efficiently and skillfully chat with the users or patients especially in Bengali users.

3 Proposed Approach

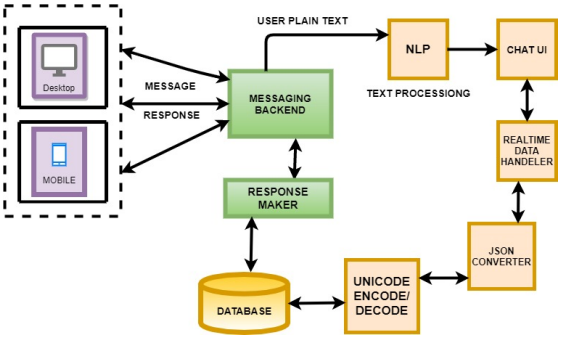

Fig. 1. Block Diagram Of the System

A detailed structure of the PRERONA: chatbot is illustrated in figure 1. The following sections discuss the different blocks related to the model. First of all, the user will type text that text will take as an input to the messaging backend, it will send to the NLP unit which is basically a text processing unit[8]. The chat user interface receives data as well. At the same time the real-time data handler catches all the data and sends it to JSON conversion which is a JavaScript text based format for presenting structured data. Now it sends the data to the UTF-8 which is a variable-width character encoding system used for electronic communication [2]. Our custom dataset receives data from UTF-8 and checks if it has that keyword or input in the dataset or not. If it matches with the dataset, sends the reply to the response makers. Otherwise it will show our default error message "Sorry not be able to understand you. Please tell me again." This is how our system responds to user messages.

3.1 Flowchart of the System

In figure 2, we can see user input is sent to the manual operation which is the user interface. Then a decision maker checks both network and internet connectivity on or off. If it finds off, it directly goes to the error handling process. This process again sends the data to the UI module so that we can say users must have an internet connection to access our application. Once again if decision makers find connectivity is okay, then it sends data to the data handler. From here, data can send to or receive from servers; databases as well as the direction flow between these are bidirectional. After that data handler process sends

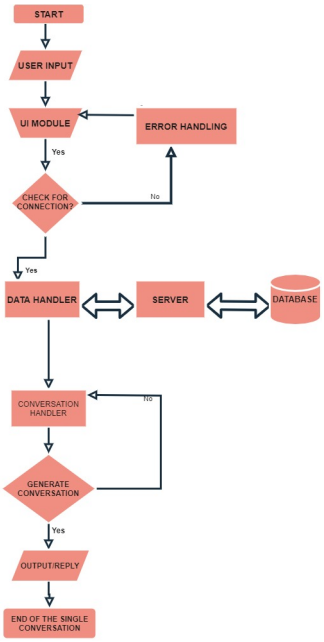

Fig. 2. Flowchart Of the System

data to the conversion handler. Again a decision maker comes, it checks whether the data generated is converted or not. If not, it sends data to the conversion handler. Otherwise, it generates data, gives output or reply as well. And then it terminates the conversation.

3.2 Algorithms

Mental Health context classifier (MHCC)

Pseudo code:

INPUT: A BUNCH OF USER TEXT
OUTPUT: DIRECTION TOWARDS SOLUTIONS

NOTATION:

TQ: Text Question

TA: Text Answer

INITIALISATION:

1. FIRST STATEMENT

   LOOP PROCESS

2. WHILE (T IN TEXT) DO
3. sp_cha = explode_speacial_character(T)
4. seg = segmentation(T)
5. Done
6. WHILE (STOP WORD IN T)
7. DELETE STOP WORD
8. GOTO STEP 3
9. DONE
10. TQ = json_data_transfer()
11. TA = find_relative_answer()
12. IF (TQ ⇔ TA) > 90% THEN
13. specific direction to user
14. JUMP A
15. ELSE IF (TQ⇔TA)>50% && < 90%
16. probabilistic direction to user
17. JUMP B
18. ELSE
19. Ask_QUERY_AGAIN()
20. A : Solution_Provided_More_Accurate
21. B: Solution_Provided_Less_Accurate+suggestions
22. END IF
23. END

3.3 Flowchart of the Algorithm

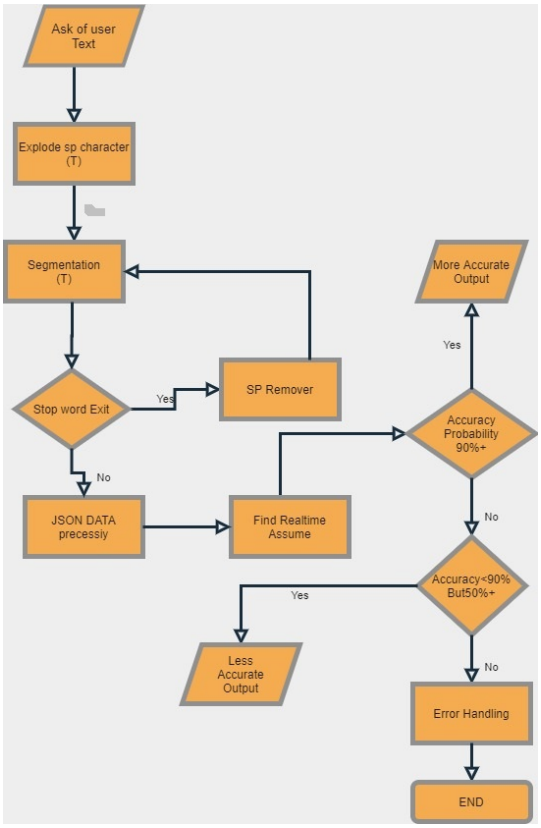

Fig. 3. Flowchart Of the Algorithm

3.4 Dataset Description

Bengali Corpus

We collected data through interviews with lots of people who were asked questions about depression. Since there is no open source Bengali mental health database for NLP, we have followed the interview questionnaire approach for collecting Bengali data. We have created a customized Bengali dataset by analyzing our collected data with the help of some mental health specialists.

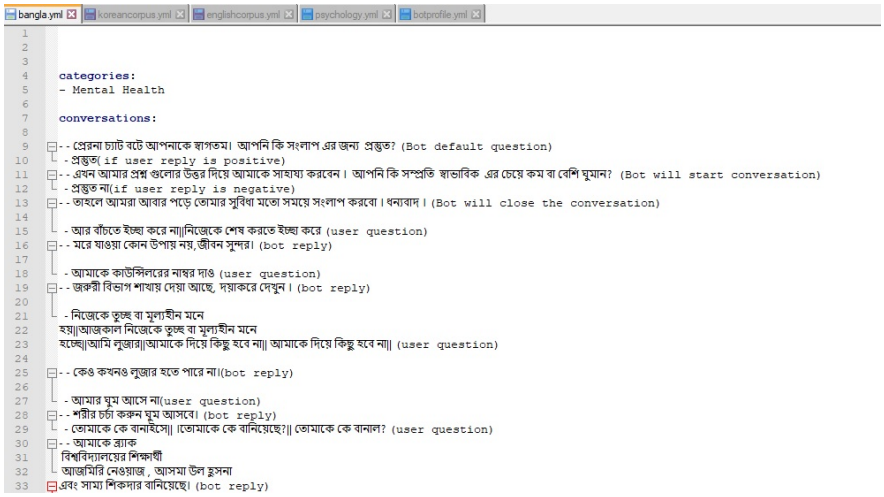

Fig. 4. A Sample Bengali Corpus

Input-Output Pattern Of our chatbot is given in figure 5. This figure is actually a sample pattern of our dataset, how we have connected similar types of input as well as it also shows our answer pattern for the predicted input.

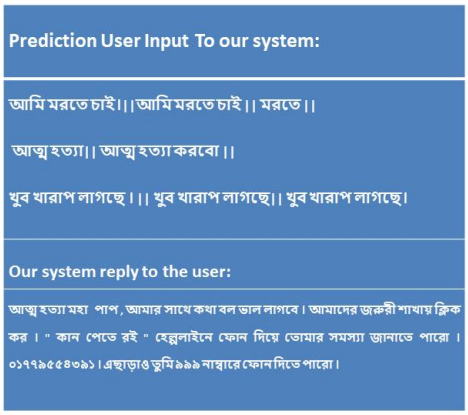

Fig. 5. Input-Output Pattern Of the System

We have allowed some users to use our system for testing purpose. In figure 6, Conversations between users and our bot is given -

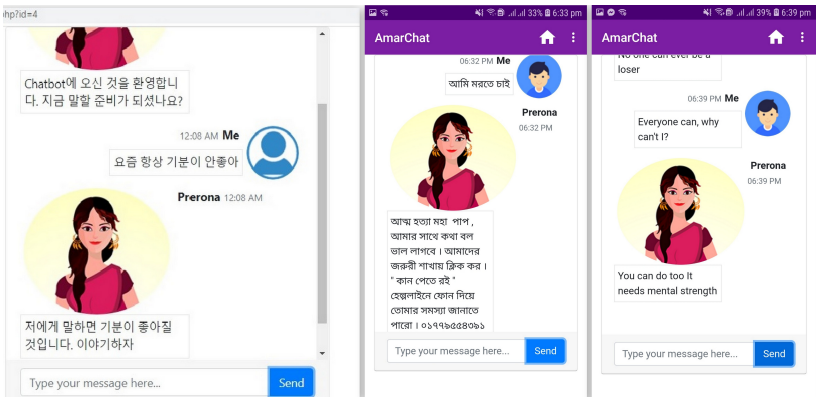

Fig. 6. Testing Data with Different Types Of the Users

4 Result Analysis

In the experiment we gave our chatbot to some users to chat. Our chatbot can chat in three different languages as a result we have got three types of results. Besides that, we have done our experiment into two phases. The accuracy of our chat depends on three things such as Database of our system, Typing skills of user, Data processing efficiency of our system. First, we have developed our Bengali database based on practical experience. The accuracy rate increases along with database enrichment. Second, it is a challenge for the user to type Bengali correctly as each letter of the Bengali alphabet has a different format. So it might be very difficult for any system to communicate with the user. We have emphasized keyword-based communication to increase performance. Third, we found our system data processing efficiency is above 98%. To mention here that users are not confined to selective questions or answers. In addition it is important for the user to type the question or the key word correctly. We have found this problem mostly in Bengali conjunctive letters.

Accuracy for Bengali Language Phase 1

In the first phase for Bengali language, our chatbot showed 50% accuracy for the two users, 45% for the two users, 40% for the seven users, 35% for the two users and it showed lowest 20% for the two users. The average accuracy for the Bengali language was 35%. From the first phase we tried to collect a lot of Bengali words and those words were added to our dataset. As a result, in the second phase we get some improvements in the accuracy.

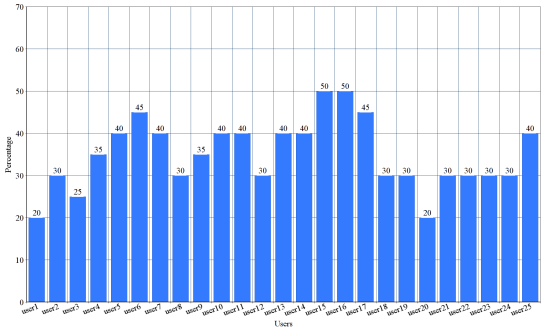

Fig. 7. Accuracy for Bengali Language Phase 1

Phase 2

From the 2nd phase, we can see the improvements in the accuracy, got highest

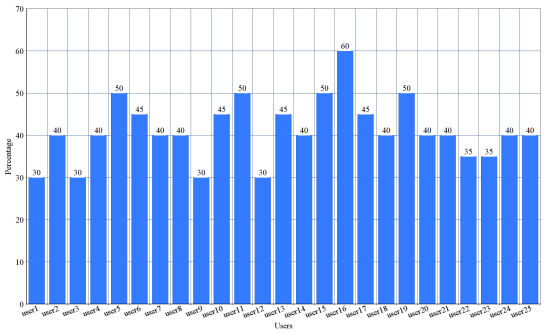

Fig. 8. Accuracy for Bengali Language Phase 2

60% accuracy from the one user, 50% accuracy from the four users, and 45% accuracy from the four users, 40% for the ten users and lowest 30% for the four users. In addition the average accuracy in the 2nd phase was boosted up to 41.2% from 35%.

Accuracy for English Language

In the case of the English language, there are no conjunctive letters. So in the English language, the accuracy rate is higher than in the Bengali language. The accuracy rate is proportional to the vastness of our database. If wrong

punctuation and spelling mistakes are reduced from user end, efficiency will increase.

Phase 1

For the 1st phase of the English language the accuracy rate was not so high. It was below average. As we can see from the figure 9, it gives the highest 50% accuracy for the one user, 45% accuracy for the three users, 40% accuracy for the seven users, 35% accuracy for the four users, and the lowest 25% accuracy for the two users. It gives us an average 35.8% accuracy.

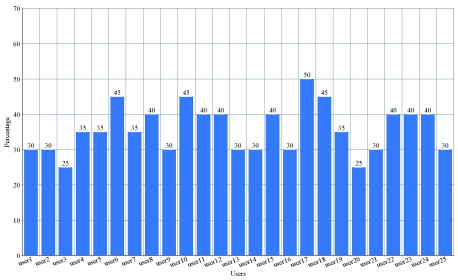

Fig. 9. Accuracy for English Language Phase 1

Phase 2

But in the 2nd phase, we can see some change in the accuracy. As we mentioned before that we use lots of words from the users in the first phase which were not included in the dataset. So there is an improvement in the accuracy for phase two from figure 10, we can see that we got the highest 50% accuracy from the four users, 45% accuracy from the six users, 40% accuracy from the eleven users, and the lowest 30% for the three users. We got 40% average accuracy in the 2nd phase which is 4.2% more than the 1st phase.

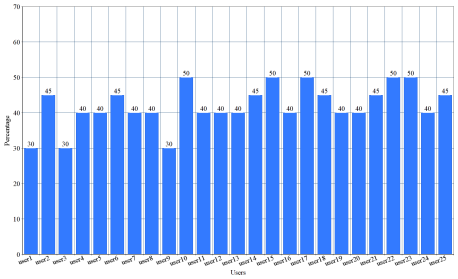

Fig. 10. Accuracy for English Language Phase 2

Accuracy for Korean Language

For the Korean language, we did our accuracy test in two phases. We could not interview native Korean people for the Korean language, initially; we have included the translated version of the English dataset. Moreover, maximum users who communicated with the bot as a test case have also used the translated method for Korean which decreased the accuracy rate. In the early phase, it gave us poor accuracy. In the first phase for the Korean language, our chatbot showed 40% accuracy for the three users, 35% for the one user, 30% for the eleven users, 20% for the eight users and it showed the lowest 10% for the one users. The average accuracy was 26.2%.

Phase 1

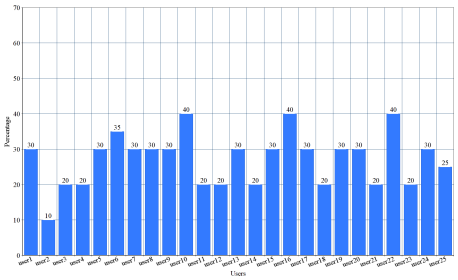

Fig. 11. Accuracy for Korean Language Phase 1

Similar to other languages we try to collect lots of Korean words and those words were added to our dataset. As a result, in the second phase, we got some

improvements in accuracy.

Phase 2

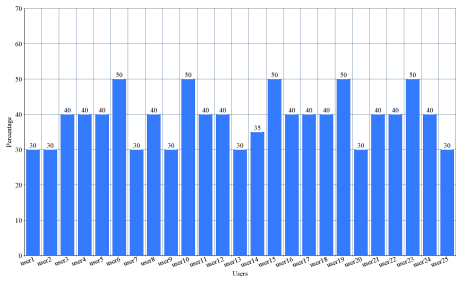

Fig. 12. Accuracy for Korean Language Phase 2

From the 2nd phase, we can notice some improvements and got highest 50% accuracy from the five users, 40% accuracy from the twelve users, and 35% accuracy from the one user and lowest 30% for the seven users. In addition, the average accuracy increases from 26.2% to 39%.

Data Accuracy

We can see from the table 1 that in the phase 1 our chatbot gave 11.61% accuracy for the Bengali language, 11.94% accuracy for the English Language and 9.07% accuracy for the Korean Language. In total it gave 32.62% accuracy and 67.38% error.

Data Accuracy Percentage for Phase 1 is shown in Table 1-

Bengali Language	English Language	Korean Language	Error
11.61%	11.94%	9.07%	67.38%

Table 1. Data Accuracy Percentage for Phase 1

But in the phase 2, we can see from the table 2 that it gave 13.70% accuracy which is 2.09% more than phase 1 for the Bengali language. Furthermore, it gave 14.00% accuracy for the English Language and 12.93% accuracy for the Korean Language from the Table 2 . In total it gave 40.63% accuracy and 59.37% error which is 8.01% lesser error than phase 1.

Data Accuracy Percentage for Phase 2 is shown in Table 2-

Bengali Language	English Language	Korean Language	Error
13.70%	14.00%	12.93%	59.37%

Table 2. Data Accuracy Percentage for Phase 2

5 Conclusion and Future Work

Our research is a pioneering work in the field of Bengali mental health chatbots. In Bangladesh, most people are not familiar with the word 'Depression', Frustration' and 'Anxiety' but they face it regularly without even realizing. Moreover, the number of suicides in Bangladesh is increasing day by day due to depression.The main goal of our software system is to provide knowledge about depression to Bengali people through our bots, talk to depressed people, and help them to overcome depression by giving positive support. The main challenge of our software system was to create a chatbot that can respond to users accurately. Due to the lack of Bengali dataset, we have implemented an supervised learning chatbot along with adaptive feature which will converse with the user based on the pattern matching algorithm. Our Bengali data corpus will help in creating a system for Bengali language processing research. In the future, we will train our bot and add machine learning as well as neural network technique. We have a plan to optimize the automated system by making the chatbot a voice-enabled system that will help to better understand people with low literacy. In addition we try to build a smart AI based chatbot which will assist users and give more accuracy. Also, we want to develop our own dataset. From our dataset we can analyze the changing in human behavior. In addition, we will add emotional intelligence techniques to the chatbot. This will further help in understanding the dialogues and identifying the user's emotions. As a result, the effectiveness of counseling will be improved.

References

[1] J. Weizenbaum, "Eliza—a computer program for the study of natural language communication between man and machine," *Communications of the ACM*, vol. 9, no. 1, pp. 36–45, 1966.

[2] F. Yergeau, "Utf-8, a transformation format of iso 10646," STD 63, RFC 3629, November, Tech. Rep., 2003.

[3] A. R. Chowdhury, A. Biswas, S. F. Hasan, T. M. Rahman, and J. Uddin, "Bengali sign language to text conversion using artificial neural network and support vector machine," in *2017 3rd International Conference on Electrical Information and Communication Technology (EICT)*, IEEE, 2017, pp. 1–4.

[4] K. K. Fitzpatrick, A. Darcy, and M. Vierhile, "Delivering cognitive behavior therapy to young adults with symptoms of depression and anxiety using a fully automated conversational agent (woebot): A randomized controlled trial," *JMIR Mental Health*, vol. 4, no. 2, 2017. DOI: 10.2196/mental.7785.

[5] D. A. Helal, "Depression: Let's talk," *The Daily Star*, Apr. 2017. [Online]. Available: https://www.thedailystar.net/health/depression-lets-talk-1384978.

[6] W. H. Organization *et al.*, "Depression and other common mental disorders: Global health estimates," World Health Organization, Tech. Rep., 2017.

[7] V. Sharma, M. Goyal, and D. Malik, "An intelligent behaviour shown by chatbot system," *International Journal of New Technology and Research*, vol. 3, no. 4, 2017.

[8] A. U. Nasib, H. Kabir, R. Ahmed, and J. Uddin, "A real time speech to text conversion technique for bengali language," in *2018 International Conference on Computer, Communication, Chemical, Material and Electronic Engineering (IC4ME2)*, IEEE, 2018, pp. 1–4.

[9] M. Kowsher, F. S. Tithi, M. A. Alam, M. N. Huda, M. M. Moheuddin, and M. G. Rosul, "Doly: Bengali chatbot for bengali education," in *2019 1st International Conference on Advances in Science, Engineering and Robotics Technology (ICASERT)*, IEEE, 2019, pp. 1–6.

[10] M. T. ZEMČÍK, "A brief history of chatbots," *DEStech Transactions on Computer Science and Engineering*, no. aicae, 2019.

[11] J. Frankenfield, *Chatbot*, Aug. 2020. [Online]. Available: https://www.investopedia.com/terms/c/chatbot.asp.

[12] *Number of people with depression increases.* [Online]. Available: https://www.who.int/bangladesh/news/detail/28-02-2017-number-of-people-with-depression-increases.

[13] *Your safe space in this difficult time.* [Online]. Available: https://www.wysa.io/.