

PRERONA- Mental Health Bengali Chatbot For Digital Counselling

by

Asma Ul Hussna

15201045

Azmiri Newaz Khan Laz

19341030

Md Shammyo Sikder

16301131

A thesis submitted to the Department of Computer Science and Engineering
in partial fulfillment of the requirements for the degree of
B.Sc. in Computer Science

Department of Computer Science and Engineering
Brac University
October 2020

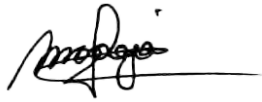
© 2020. Brac University
All rights reserved.

Declaration

It is hereby declared that

1. The thesis submitted is my/our own original work while completing degree at Brac University.
2. The thesis does not contain material previously published or written by a third party, except where this is appropriately cited through full and accurate referencing.
3. The thesis does not contain material which has been accepted, or submitted, for any other degree or diploma at a university or other institution.
4. We have acknowledged all main sources of help.

Student's Full Name & Signature:



Asma Ul Hussna
15201045



Azmiri Newaz Khan Laz
19341030



Md. Shammyo Sikder
16301131

Approval

The thesis/project titled “PRERONA- Mental Health Bengali Chatbot For Digital Counselling ” submitted by

1. Asma Ul Hussna (15201045)
2. Azmiri Newaz Khan Laz (19341030)
3. Md. Shammyo Sikder (16301131)

Of Summer, 2020 has been accepted as satisfactory in partial fulfillment of the requirement for the degree of B.Sc. in Computer Science on October 9, 2020.

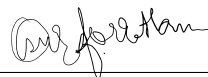
Examining Committee:

Supervisor:
(Member)



Dr. Jia Uddin
Assistant Professor(Research)
Technology Studies
Woosong University
Associate Professor(On leave)
Department of Computer Science and Engineering
Brac University

Co-supervisor:
(Member)



A.M. Esfar-E- Alam
Lecturer
Department of Computer Science and Engineering
Brac University

Head of Department:
(Chair)

Prof. Mahbub Majumdar
Chairperson
Dept. of Computer Science & Engineering
Brac University

Dr. Mahbubul Alam Majumder
Professor
Department of Computer Science and Engineering
Brac University

Ethics Statement

The thesis is carried out in complete compliance with research ethics norms, and the codes and practices set by BRAC University. In our thesis we use the data from primary sources. We collect data from group interview and we use our own data set for our thesis. We are ensuring we use references and in text citations properly. We the three co-authors take full responsibility for the thesis code violations. For solving problems we read lots of websites, YouTube tutorials, and Questionnaire Free tools etc. We also took help from our university faculty members. Finally, We declare that we give credit every people from whom we took help. We did not make any fraud-able means for completing the thesis. Our work is in compliance with the ethics standard set by BRAC University.

Abstract

"Human", "Machine", "Interaction" this word refers to the modern era of technology where researchers give emphasis on the interaction between computer machines and humans. We humans cannot confine ourselves depending on only human-based health care systems. The current time of digitalization brings with it progress and additional opportunities for medical care in general and clinical psychology in particular. In this paper, we have presented a model named PRERONA: a psychological Bengali chatbot for depressed persons. PRERONA is designed to help depression struggling especially for Bengali users who are lack attention or do not have anyone to talk about their problem. The main goal of this chatbot "Prerona" is to give instant answers to questions and queries as well as to give proper mental health to Bengalis in times of depression. It is able to carry out real-time chats in initially three languages, Bangla, English, and Korean. Natural language processing (NLP) is used here to successfully implement these languages. First, it matches the input with the query that exists in the database, then it identifies each matching sentence with the input, and finally, it selects the one with the highest suite in response to the input. To test the proposed chatbot, we have interviewed 300 persons to come up with the datasets. Existing chatbots such as Woebot, Wysa, Moodkit, etc. do not allow users to express them. However, PRERONA allows the user space to talk about whatever they want and it is intelligent enough to ask questions as well as answer questions.

Keywords: NLP; Interaction; Psychology; Supervised; Adaptive; Anxiety; Health care; Chatbot; Depression

Dedication

We would like to dedicate this research to our parents. Without their support we may not be able to do our studies. We also want to dedicate the research to our friends who helped us to do better and most importantly the participants who helped us collecting data from giving us an interview. Both our supervisor and co-supervisor helped us throughout the year. We also want to dedicate our paper to them.

Acknowledgement

First and foremost, we are grateful to God the Almighty, to whom the wisdom and power belongs, for His blessings and love to His creations.

Secondly, our appreciation goes to our thesis supervisor Dr. Jia Uddin and our co-supervisor A.M.Esfar-E-Alam for granting a wonderful opportunity to work with them and to support and guide us in this thesis.

Thirdly, We are also indebted to everyone who helped us to proceed in our thesis work.

And finally to our parents without their throughout support it may not be possible. With their kind support and prayer we are now on the verge of our graduation.

Table of Contents

Declaration	i
Approval	ii
Ethics Statement	iii
Abstract	iv
Dedication	v
Acknowledgment	vi
Table of Contents	vii
List of Figures	ix
List of Tables	x
Nomenclature	xi
1 Introduction	1
1.1 Motivation	2
1.2 Problem Statements	2
1.3 Objectives and Contributions	3
1.4 Thesis Orientation	4
2 Background	5
2.1 Literature Review	5
2.2 Related Works of Our Research	5
3 Proposed Approach	6
3.1 Block Diagram of Our System PRERONA	6
3.2 Workflow of Our System	7
3.3 Flowchart of Our System	7
3.4 Use Case Diagram Of Our System	8
3.5 Architecture Model of PRERONA Bot	10
3.6 Software Development Method	10

4	Algorithm and Data Description	11
4.1	Algorithm	11
4.2	Algorithm Flowchart	12
4.3	Dataset Description	13
4.3.1	Bengali Dataset	13
4.3.2	English Dataset	14
4.3.3	Korean Dataset	15
4.4	Testing Chatbot-PRERONA with Different Types of The Users . . .	16
5	Result Analysis	20
5.1	Accuracy for Bengali Language	20
5.2	Accuracy for English Language	22
5.3	Accuracy for Korean Language	24
5.4	Comparison Result Between Phase 1 and Phase 2	26
6	Conclusion and Future Work	28
6.1	Conclusion	28
6.2	Future Work	28
	Bibliography	31

List of Figures

1.1	Cases of Depression Disorder	2
1.2	Suicide Cases In Bangladesh	3
3.1	Block Diagram of Our System PRERONA	6
3.2	Server and User Interface - Workflow Diagram	7
3.3	Flowchart of Our System	8
3.4	Use Case Diagram	9
3.5	Retrieval-Based Model	10
3.6	Spiral Model	10
4.1	Algorithm Flowchart	12
4.2	A Sample Bengali Corpus	13
4.3	Input-Output Pattern of the System	14
4.4	A Sample English Corpus	15
4.5	A Sample Korean Corpus	16
4.6	Bengali User Conversation with Our Bot	17
4.7	English User Conversation with Our Bot	18
4.8	Korean User Conversation with Our Bot	19
5.1	Accuracy for Bengali Language (1st Phase)	21
5.2	Accuracy for Bengali Language (2nd Phase)	22
5.3	Accuracy for English Language (1st Phase)	23
5.4	Accuracy for English Language (2nd Phase)	24
5.5	Accuracy for Korean Language (1st Phase)	25
5.6	Accuracy for Korean Language (2nd Phase)	26
5.7	Comparison Between Phase 1 and Phase 2	27

List of Tables

5.1	Data Accuracy Percentage for Phase 1	27
5.2	Data Accuracy Percentage for Phase 2	27

Nomenclature

The next list describes several symbols & abbreviation that will be later used within the body of the document

API Application Programming Interface

CBI Cognitive Behavioral Therapy

HTML Hyper Text Markup Language

JSON JavaScript Object Notation

MIT Massachusetts Institute of Technology

NLP Natural Language Processing

UFT8 8- bit Unicode Transformation Format

UI User interface

WHO World Health Organisation

Chapter 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 [15] [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 [3]. 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%) [17]. 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 [4]. But this is even more shocking that In spite of the existing treatments for mental disorders, 76%-86% people of low to middle income country do not seek medical attention for their disorders either out of negligence or hesitation due to social stigma. Here mental health chatbot plays the most vital role. Mental health chatbots are specifically integrated so that they can have conversation with the patient whom we consider to be a user about their mental health and wellbeing, in addition to that the patient will find the bot beside them 24/7. Many such chatbots namely Weobot, Wysa were earlier designed to help depression struggling patients [16]. Both of these follow cognitive behavioral therapy (CBT) which is evidence based approach to treatment but these bots seemed a bit scripted and do not give their users access to share their thoughts which make it more like an interactive quiz rather than a chat [9]. Besides there is Dolly and Golpo which are Bengali chatbots but used in different sectors. Here PRERONA: mental health chatbot for depressed people is updated. Besides asking real time questions to determine and fix the mental condition of the user, PRERONA lets its users share their feelings and also responses accordingly. Additionally it provides the users with lot of contents relating anti-depression and anti-anxiety[9]. PRERONA is not confined to Bangladesh only. It is a natural language processing (NLP) based tri-language chatbot where users can chat in Bengali, English and Korean which makes it more interesting [13]. The core objective of this bot is to bring more glamorous variation in mental health chatbot by linking to mindfulness rather than being straight forward. This paper presents a detailed analysis of the proposed model of Prerona, experimental setups and result analysis.

1.1 Motivation

Chatbots occupy the central role as computer-human interfaces. This is a kind of software which makes humans think that they are talking to another human by simulating typed conversation. Using Natural Language Processing (NLP) we have created a conversational agent which we call chatbot, can talk to any user regardless what the field is [13]. In our digital and automated world, humans might often keep their feelings unexpressed which lead them to many mental diseases. 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]. Work on mental health chatbots is still new in Bangladesh and we have not found any work on this topic especially for Bangladesh. There are a few Bengali chatbots for example Dolly and Golpo but they do not support mental health systems. Internationally, online mental health is a briskly developed substitute for orthodox medical consultation and with the rapid escalation in need of supplementary support, the bot targets to make an ambience that is more conversational than automated for users in need[9].

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.

1.2 Problem Statements

Depression has become an alarming problem world-wide. Statistics shows that depression cases are highest in southeast Asia region compared to the other parts of the world, which is 27%. Besides western pacific region has 21%, American region 15%, European region 12% and Africa 9% cases of depression.

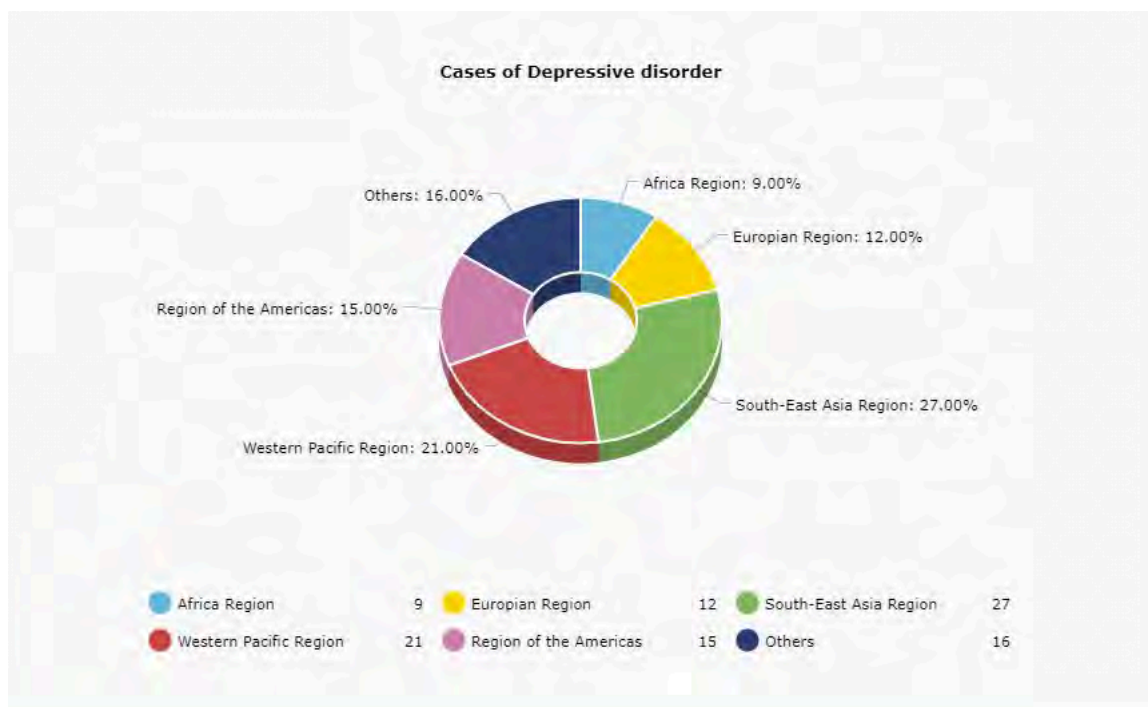


Figure 1.1: Cases of Depression Disorder

Initially, we are focusing on the mass population of Bangladesh. Bangladeshis aged 15 and over can read and write, the literacy rate in the country is about 60%-62% within this 4.1% and 4.4% of people face depression and anxiety disorder. It is even more shocking that over 53900 people has committed suicide within year 2014 -2019 [10].

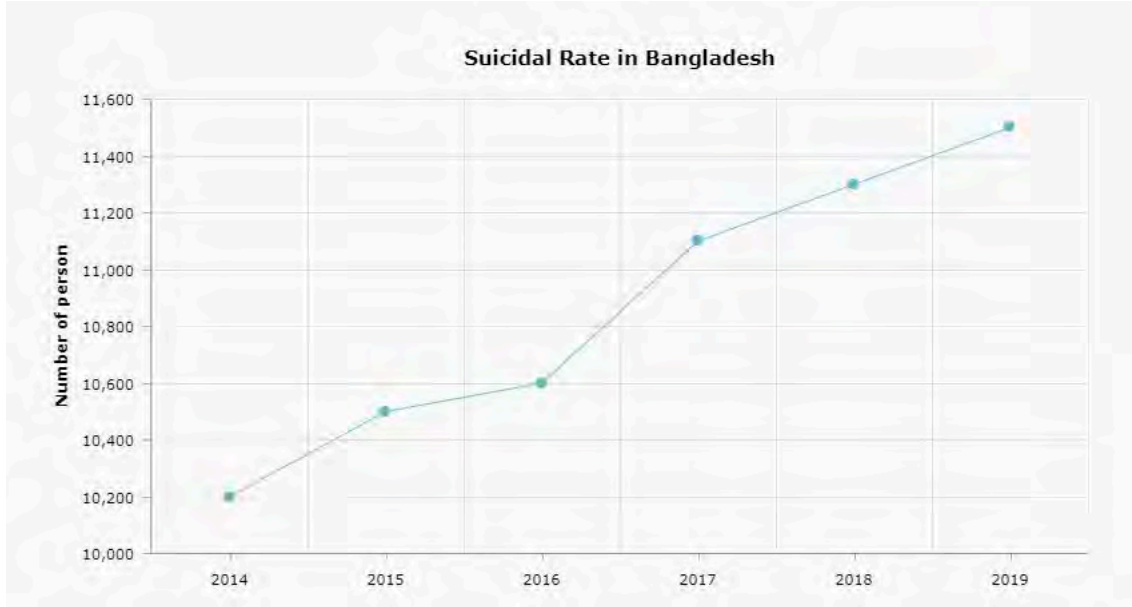


Figure 1.2: Suicide Cases In Bangladesh

To solve this major problem PRERONA serves as the Bengali language based mental health chatbot which can be accessible to any user breaking the literacy barrier. And also offers two more languages along with Bengali [8]. Due to social taboo and hesitation many keep their disease untreated, they fear to go to a psychiatrist or counselor. As a solution to this bot PRERONA provides its users or patients real-time chat to get proper guidance to conduct their mental health supervised by professionals. Moreover, it has many contents available in systems such as motivational stories, relaxing songs, books, yoga videos, etc. which will help the users to develop both their mental and physical health.

1.3 Objectives and Contributions

1. To create a chatbot in Bengali that can take input from users and then also get answers in Bengali.
2. To establish a Bengali corpus for the chatbot based on interviewing mass people under the supervision of professionals and preparing the system to bring out an enriched Bengali corpus for future use using its adaptive feature.
3. To provide the users proper guidance to heal their mental health.

1.4 Thesis Orientation

Chapter 1: Introduction where motivation, problem statement, objectives and contributions are discussed

Chapter 2: Background where literature review and related works of our research are discussed.

Chapter 3: Proposed Model, block diagram of our system, workflow between server and user, use case diagram, flowchart, architecture model, and how we developed our software described this section in detail.

Chapter 4: Data Set Description and Result analysis where user experienced analysis, software accuracy is described.

Chapter 5: Result Analysis and Data comparison.

Chapter 6: Conclusion, we conclude our research based on theory with our models that is made and discuss future works what we will do for further improvement and lastly References.

Chapter 2

Background

2.1 Literature Review

Eliza was created in an AI laboratory in MIT between 1964-66 is the oldest and well known chatbots in the history of chatbots. Its present implementation is on the MAC time-sharing system at MIT. It is written in MAD-SLIP for the IBM 7094. The man behind Eliza was professor Joseph Weizenbaum. 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] [14]. Parry is another well-known work done by Kenneth Mark Colby, was the first to pass the Turing Test - it was in the early seventies, when human interrogators, interacting with the program via remote keyboard, were unable with more than random accuracy to distinguish Parry from an actual paranoid individual. Now a days we can find a number of chatbots in practice. Alexa, SIRI has brought revolutionary change in this field. Using advanced conventional AI they can voice chat with the user and can even do tasks for them.

2.2 Related Works of Our Research

There are lots of chatbot in different sectors. Advanced technology have allowed for using of chat bots in the creation of software-based therapy. Now the chat bot will see you instead of a therapist. Early work on mental health chatbot, Woebot claims to be AI powered, follows cognitive behavioral therapy (CBT). Initially it was founded for young adults in college and graduate school. It provides an automated conversational environment but users find it a bit scripted as the queries and responses are set, users get less chance to express them. Another similar work on this field is Wysa [18]. 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 [11]. 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 user or patient especially in Bengali. As it is NLP based, it is able to understand all the language based on the database provided. Initially, we have focused on Bengali along with English and Korean languages.

Chapter 3

Proposed Approach

3.1 Block Diagram of Our System PRERONA

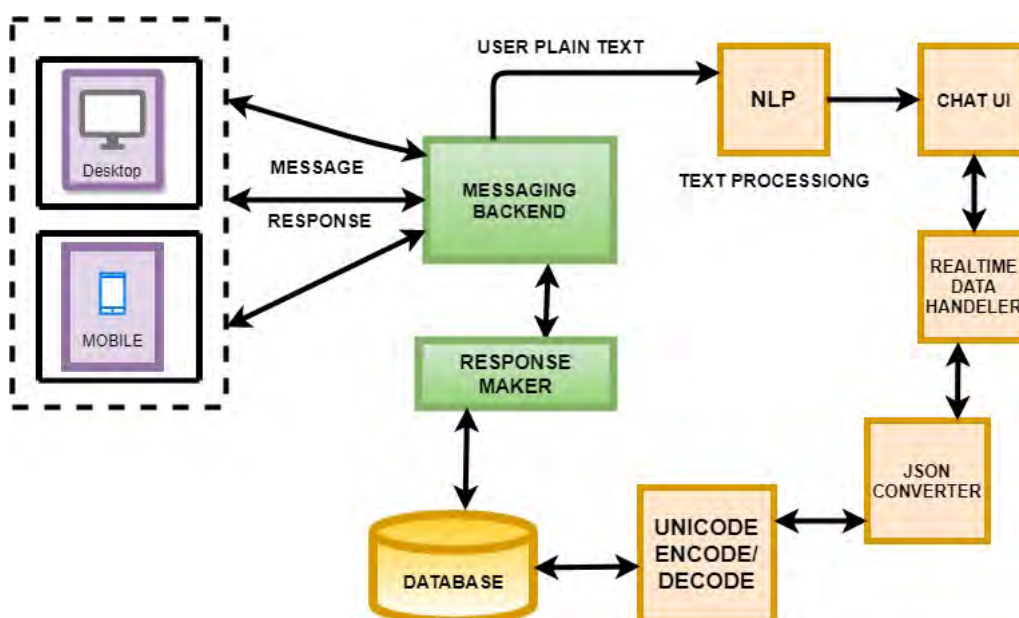


Figure 3.1: Block Diagram of Our System PRERONA

A detailed structure of the PRERONA-chatbot is illustrated in figure 3.1. 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. Chat user interface receives data as well. At the same time the real data handler catches 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 database 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.

3.2 Workflow of Our System

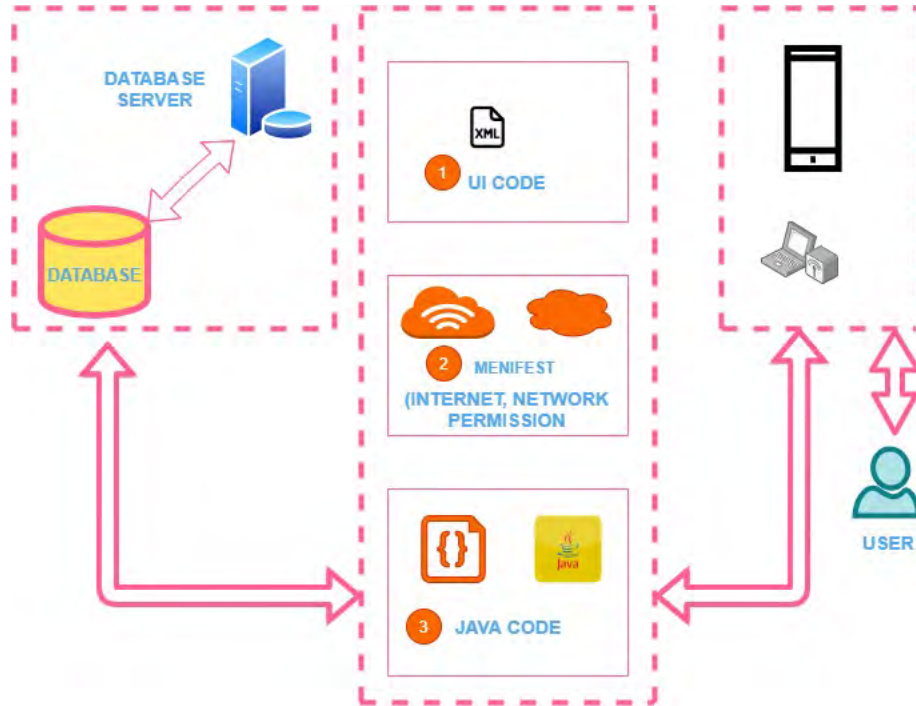


Figure 3.2: Server and User Interface - Workflow Diagram

The overall workflow of our system is given in figure 3.2. Both web and android users can access our system. Our local database is uploaded to our database server. When a user gives input text, at first the web API code generates automatically. Secondly, manifest will check the user's internet and network connection. Without internet connection users will not be able to access our system. Thirdly, java code is for user interface. It does important work such as it can call web functions, handles cookies, handle internet connectivity, from our application URL as well as it can call API data. In addition, databases are connected bidirectionally with server and android codes.

3.3 Flowchart of Our System

The flowchart diagram is shown in the figure 3.3. Here 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 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.

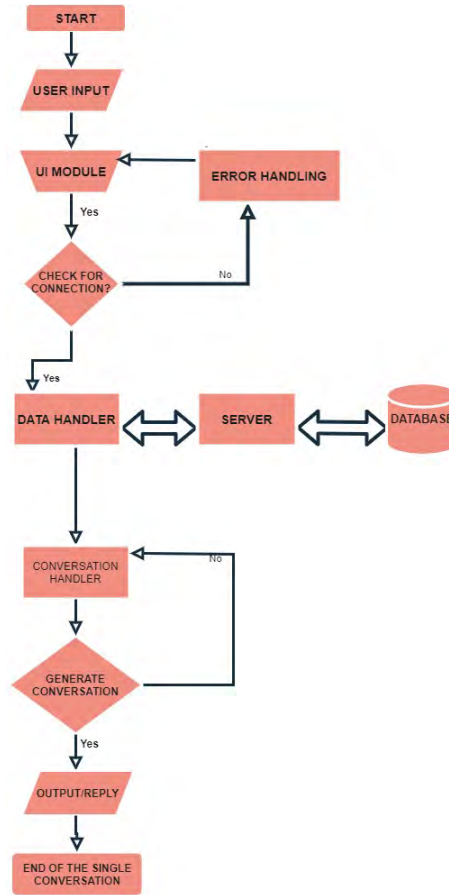


Figure 3.3: Flowchart of Our System

3.4 Use Case Diagram Of Our System

Use Case diagram of our system chatbot “PRERONA” is given in figure 3.4. We assume that only depressed users will use our application. From our use case diagram we can see the interaction between human and system. After accessing our application, a depressed user can view three types of ChatBox such as Bengali ChatBox, English ChatBox and Korean ChatBox. He will click the chat language of his choice which he wants to chat. Our main target is to give a depressed user to a positive vibe. For that we have designed our system in the following sections. Users can see other sections on the home page named Music, Exercise, Movies, Admin Advice, Online Article, Books, Motivation stories and emergency section. Each section has subsections for example in the Admin Advice section; we try to give some suggestions which will help depressed users. Users can watch some inspirational speeches taken from YouTube by clicking motivational stories. We have added some online articles that are related to awareness and to overcome depression. By reading this, users will gain knowledge of depression. In addition, our application has an emergency section where users can get some emergency numbers to contact counselors. Besides that, in our system we have added relaxation music, exercise such as yoga which will give users mental peace. By adding movies and books, we try to keep users busy, sometimes it helps users to forget simple problems. On the other hand, only administrators can access and edit datasets.

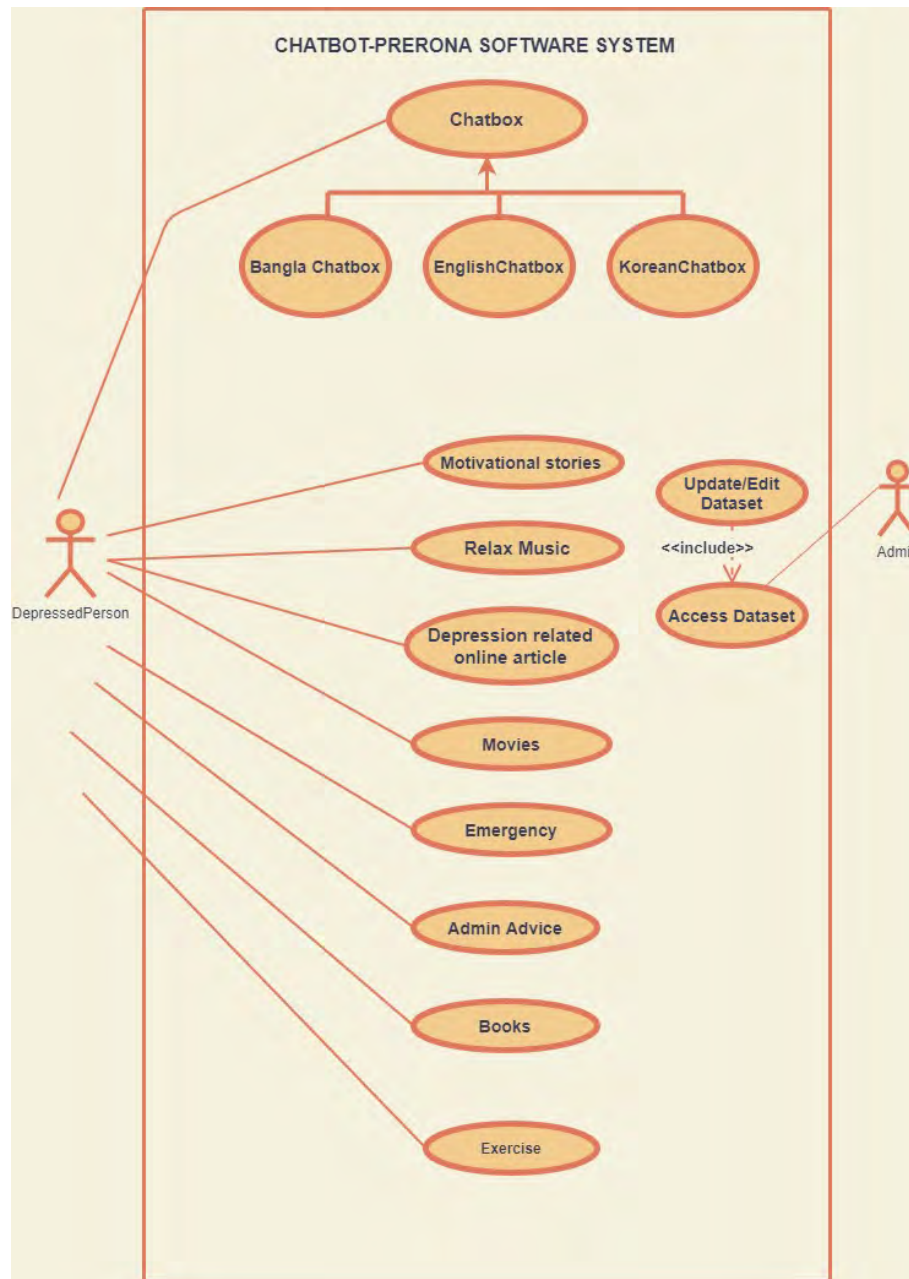


Figure 3.4: Use Case Diagram

3.5 Architecture Model of PRERONA Bot

We have used Retrieval-based models. To build this chat bot on this architecture model we used the above algorithm. This model of architecture is easy to build as well as reliable that why we choose this model.

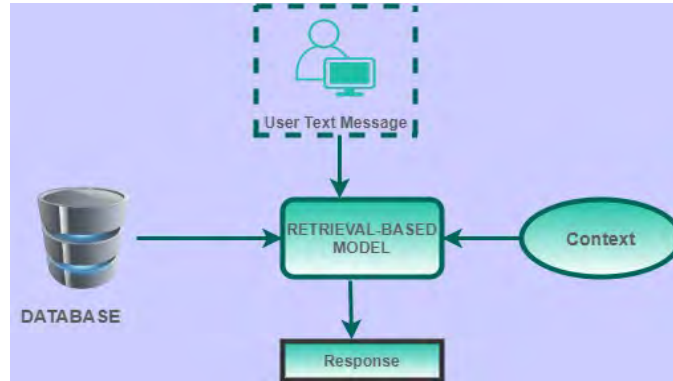


Figure 3.5: Retrieval-Based Model

3.6 Software Development Method

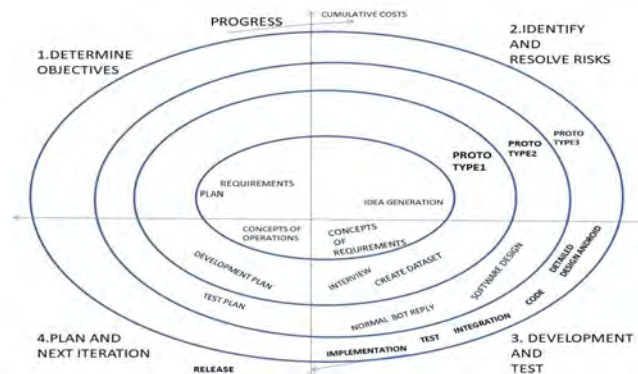


Figure 3.6: Spiral Model

We have used a spiral model to develop our software chatbot system which is shown in figure 3.6. Spiral model works based on proto-type. In prototype 1, first we generate ideas, gather requirements, create a custom dataset by interviewing and make plans for development. In prototype 2, we work on our system diagrams, normal code implementation as well as test plan. We have got a detailed android design model in prototype 3. We have implemented complex codes, integrated, tested and finally we released our chatbot application. In future we will release an updated version of the system.

Chapter 4

Algorithm and Data Description

4.1 Algorithm

Mental Health Context Classifier (MHCC)

Pseudo code:

```
Input : A bunch of user text
Output : Direction towards solutions
Notation : AF : positive answer
NG : negative answer
TQ : text question
TA : text Answer
while(text) do
  explodespchar(text)
  segmentation(T)
  while(stopwords in T)
    delstopword(T)
    TQ : jsondatatransfer(T)
    TA : findrelativeanswer(TQ)
    if(TQ <=> TA) > 90%
      specificdirectiontouser jumpA if(TQ <=> TA) > 50% & < 90%
      probabilisticdirectiontouser
      jumpB
    else
      askqueryagain
      A : solutionprovidedmoreaccurate
      B : solutionprovidedlessaccurate + withsuggestion
  END
```

This algorithm is not specifically Naive Bayes but related and modified by naive Bayes algorithms.

4.2 Algorithm Flowchart

Flowchart Algorithm of our system is given in figure 4.1

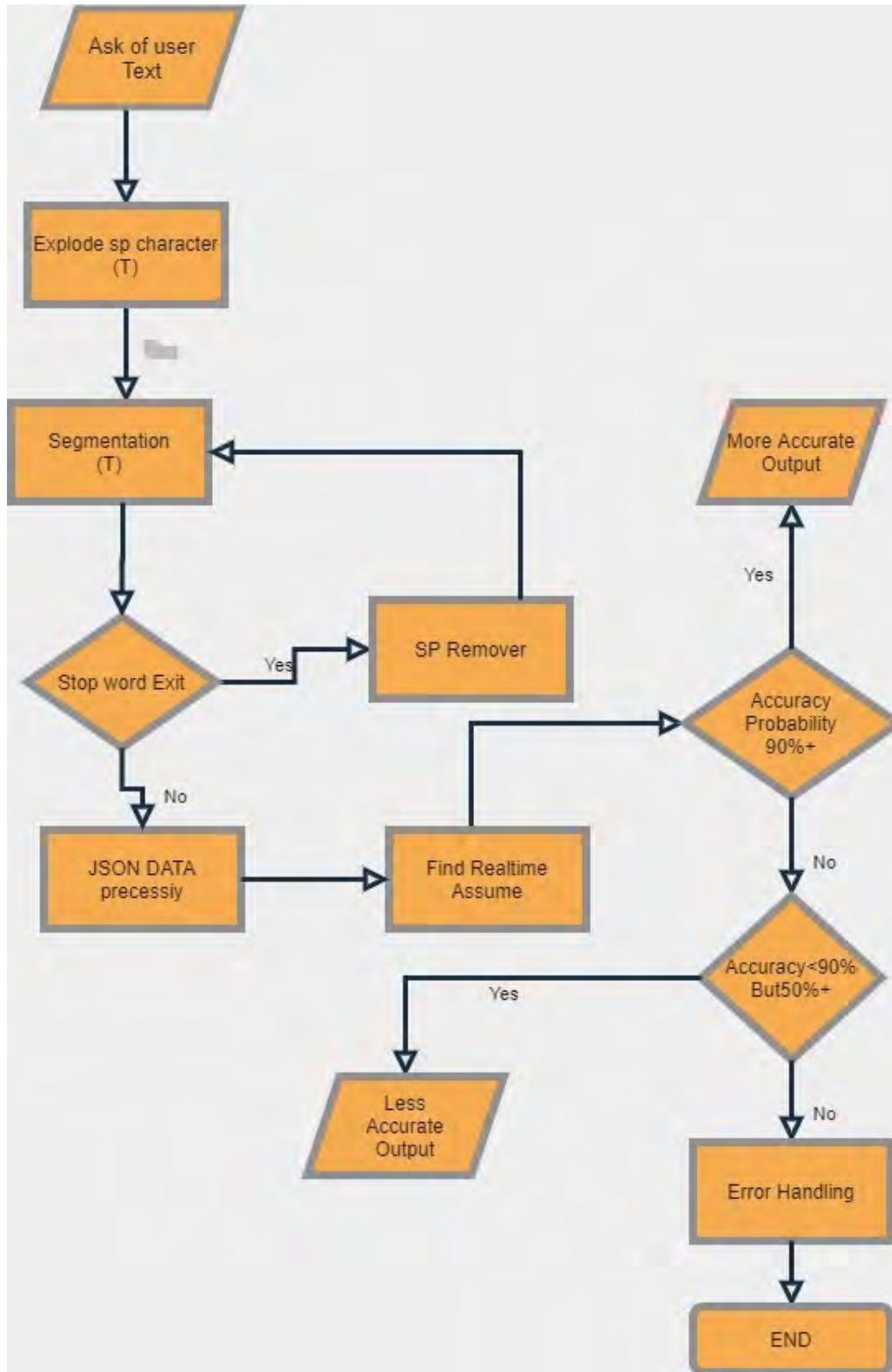


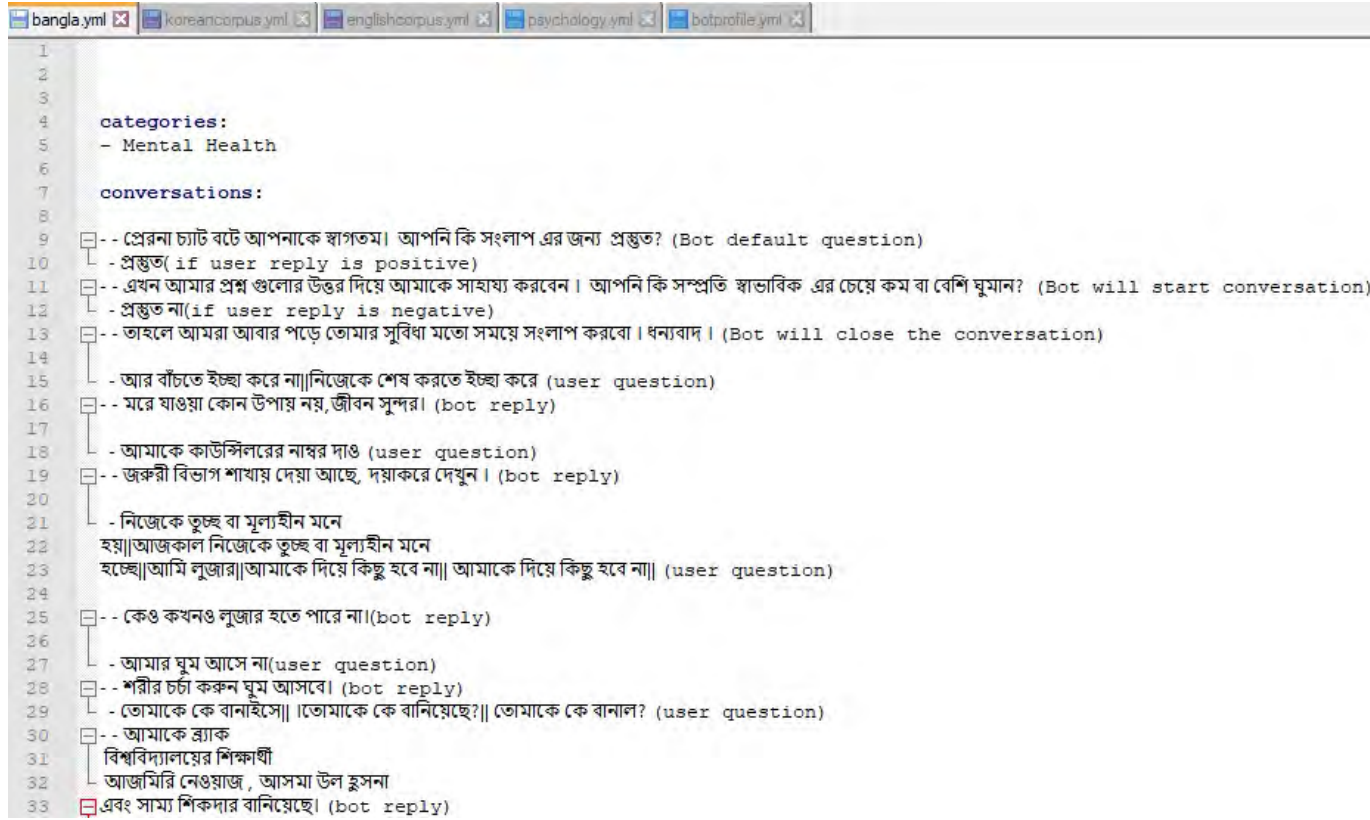
Figure 4.1: Algorithm Flowchart

4.3 Dataset Description

We collected data through interviews with 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 analyzed our collected data and created a customized Bengali data set. Although our bot preronas is made only for Bengali users, we have it in multiple languages. In addition to that, we are using adaptive learning process. If the bot can not answer any questions, that will be recorded and will be added to that dataset later on for more accuracy.

4.3.1 Bengali Dataset

In figure 4.2, A sample Bengali corpus is given-



```
1
2
3
4 categories:
5   - Mental Health
6
7 conversations:
8
9   -- প্রেরনা চ্যাট বটে আপনাকে স্বাগতম। আপনি কি সংলাপ এর জন্য প্রস্তুত? (Bot default question)
10  -- প্রস্তুত (if user reply is positive)
11  -- এখন আমার প্রশ্ন গুলোর উত্তর দিয়ে আমাকে সাহায্য করবেন। আপনি কি সম্প্রতি স্বাভাবিক এর চেয়ে কম বা বেশি ঘুমান? (Bot will start conversation)
12  -- প্রস্তুত না (if user reply is negative)
13  -- তাহলে আমরা আবার পড়ে তোমার সুবিধা মতো সময়ে সংলাপ করবো। ধন্যবাদ। (Bot will close the conversation)
14
15  -- আর বাঁচতে ইচ্ছা করে না||নিজেকে শেষ করতে ইচ্ছা করে (user question)
16  -- মরে যাওয়া কোন উপায় নয়, জীবন সুন্দর। (bot reply)
17
18  -- আমাকে কাউন্সিলরের নামের দাও (user question)
19  -- জরুরী বিভাগ শাখায় দেয়া আছে, দয়াকরে দেখুন। (bot reply)
20
21  -- নিজেকে তুচ্ছ বা মূল্যহীন মনে
22  হয়||আজকাল নিজেকে তুচ্ছ বা মূল্যহীন মনে
23  হচ্ছে||আমি লুজার||আমাকে দিয়ে কিছু হবে না|| আমাকে দিয়ে কিছু হবে না|| (user question)
24
25  -- কেও কখনও লুজার হতে পারে না।(bot reply)
26
27  -- আমার ঘুম আসে না(user question)
28  -- শরীর চর্চা করুন ঘুম আসবে। (bot reply)
29  -- তোমাকে কে বানাইসে|| তোমাকে কে বানিয়েছে?|| তোমাকে কে বানাল? (user question)
30  -- আমাকে ব্র্যাক
31  বিশ্ববিদ্যালয়ের শিক্ষার্থী
32  আজমিরি নেওয়াজ, আসমা উল হুসনা
33  -- এবং সাম্য শিকদার বানিয়েছে। (bot reply)
```

Figure 4.2: A Sample Bengali Corpus

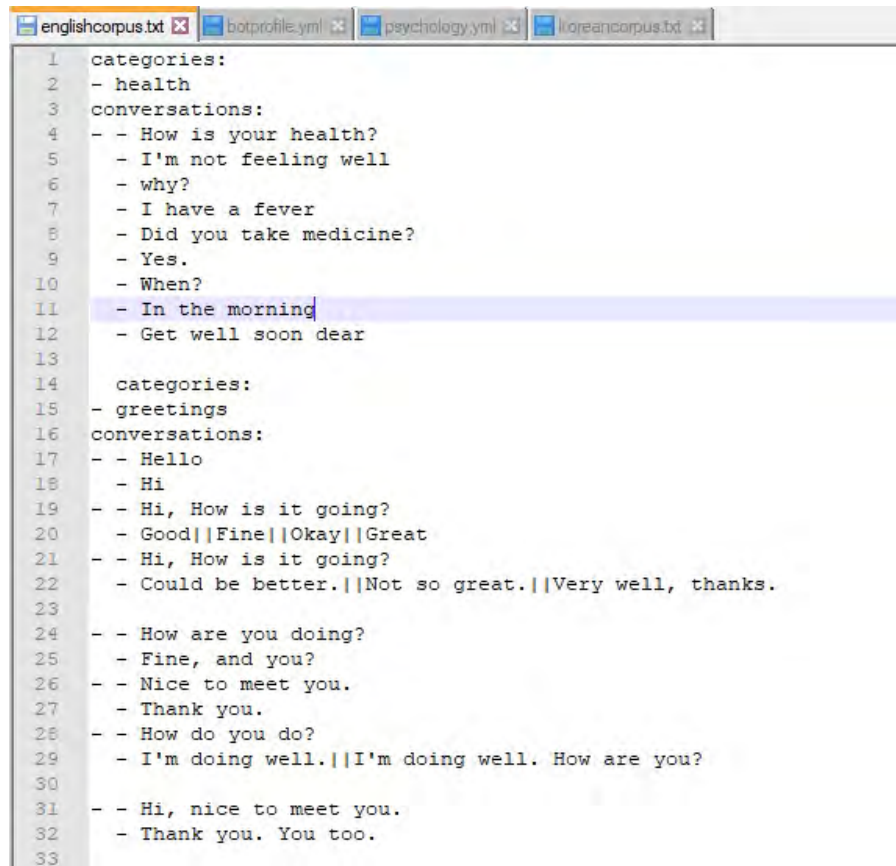
Input-Output Pattern:

Prediction User Input To our system:
আমি মরতে চাই।।। আমি মরতে চাই।। মরতে।।
আত্ম হত্যা।। আত্ম হত্যা করবো।।
খুব খারাপ লাগছে।।। খুব খারাপ লাগছে।। খুব খারাপ লাগছে।
Our system reply to the user:
আত্ম হত্যা মহা পাপ, আমার সাথে কথা বল ভাল লাগবে। আমাদের জরুরী শাখায় ক্লিক কর। " কান পেতে রই " হেল্পলাইনে ফোন দিয়ে তোমার সমস্যা জানাতে পারো। ০১৭৭৯৫৫৪৩৯১। এছাড়াও তুমি ৯৯৯ নাম্বারে ফোন দিতে পারো।

Figure 4.3: Input-Output Pattern of the System

4.3.2 English Dataset

Lots of datasets are available for English language. We have downloaded a trained data set from the Kaggle site. Also translated Bengali data sets and set it to our database. Next, we set this corpus in our dataset. A sample of English Dataset of our system is given below in figure 4.4.



```
1 categories:
2 - health
3 conversations:
4 - - How is your health?
5 - - I'm not feeling well
6 - - why?
7 - - I have a fever
8 - - Did you take medicine?
9 - - Yes.
10 - - When?
11 - - In the morning
12 - - Get well soon dear
13
14 categories:
15 - greetings
16 conversations:
17 - - Hello
18 - - Hi
19 - - Hi, How is it going?
20 - - Good||Fine||Okay||Great
21 - - Hi, How is it going?
22 - - Could be better.||Not so great.||Very well, thanks.
23
24 - - How are you doing?
25 - - Fine, and you?
26 - - Nice to meet you.
27 - - Thank you.
28 - - How do you do?
29 - - I'm doing well.||I'm doing well. How are you?
30
31 - - Hi, nice to meet you.
32 - - Thank you. You too.
33
```

Figure 4.4: A Sample English Corpus

4.3.3 Korean Dataset

Korean dataset is similar to English dataset. We have translated English and Bengali data to create Korean dataset corpus. A sample of Korean dataset corpus is shown in figure 4.5.

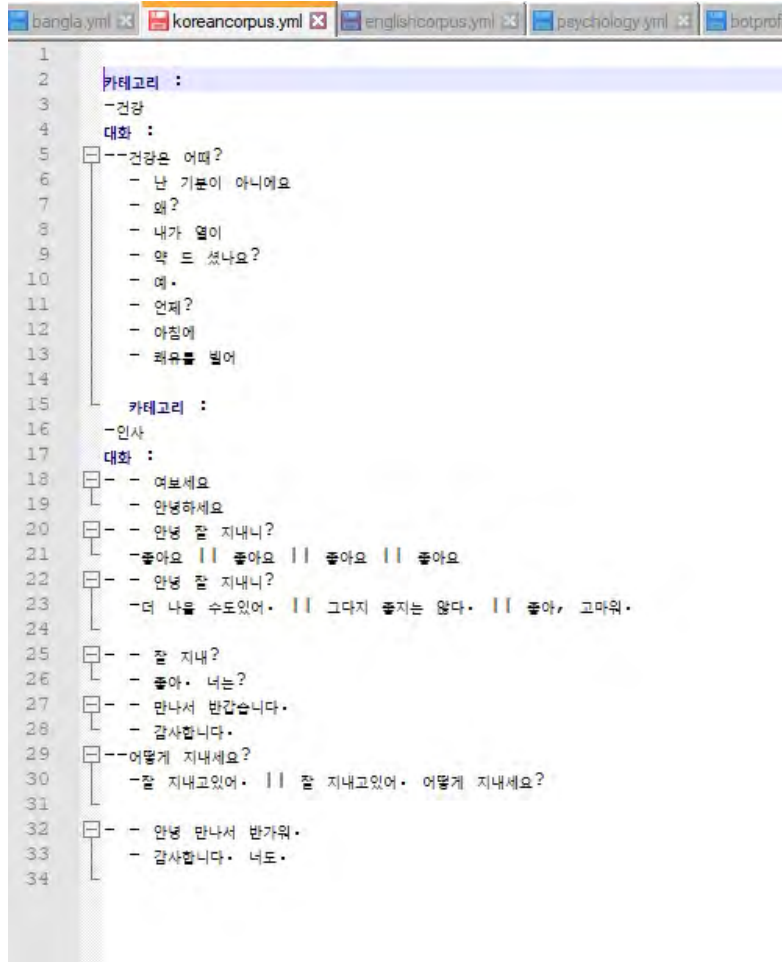


Figure 4.5: A Sample Korean Corpus

4.4 Testing Chatbot-PRERONA with Different Types of The Users

We have allowed some users to use our system for testing purpose. Conversations between users and our bot is given in figure 4.6, 4.7 and 4.8.

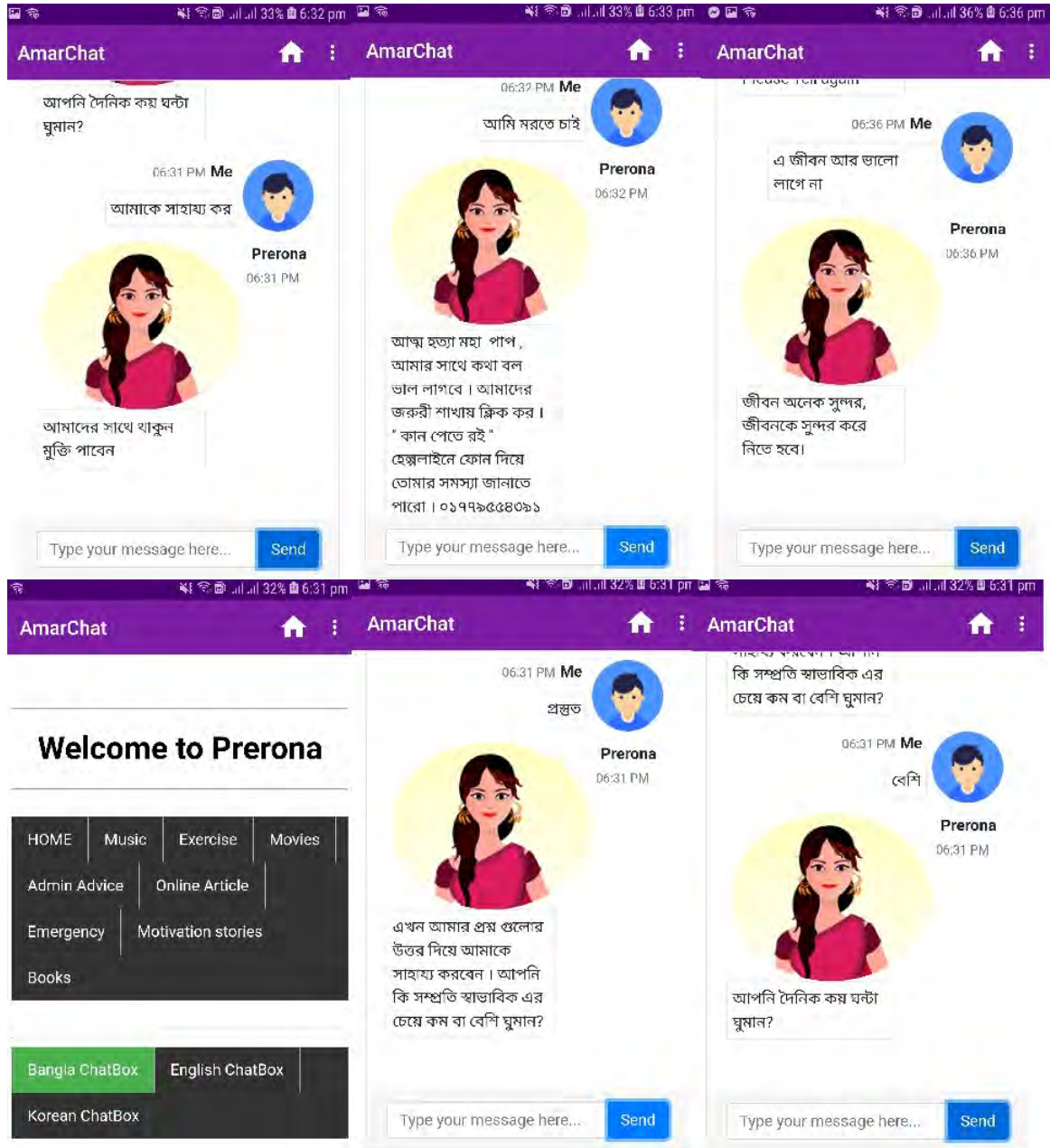


Figure 4.6: Bengali User Conversation with Our Bot

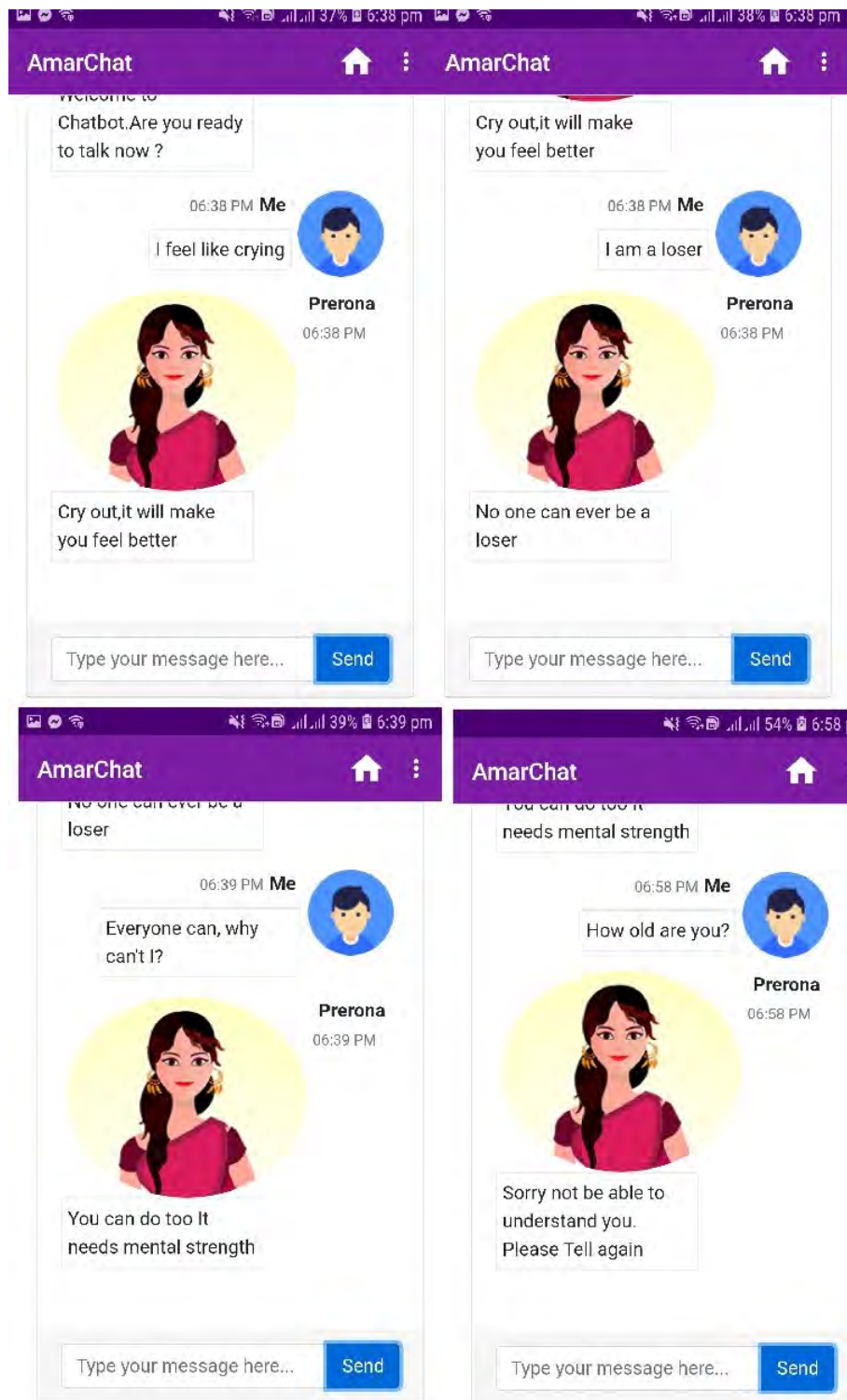


Figure 4.7: English User Conversation with Our Bot

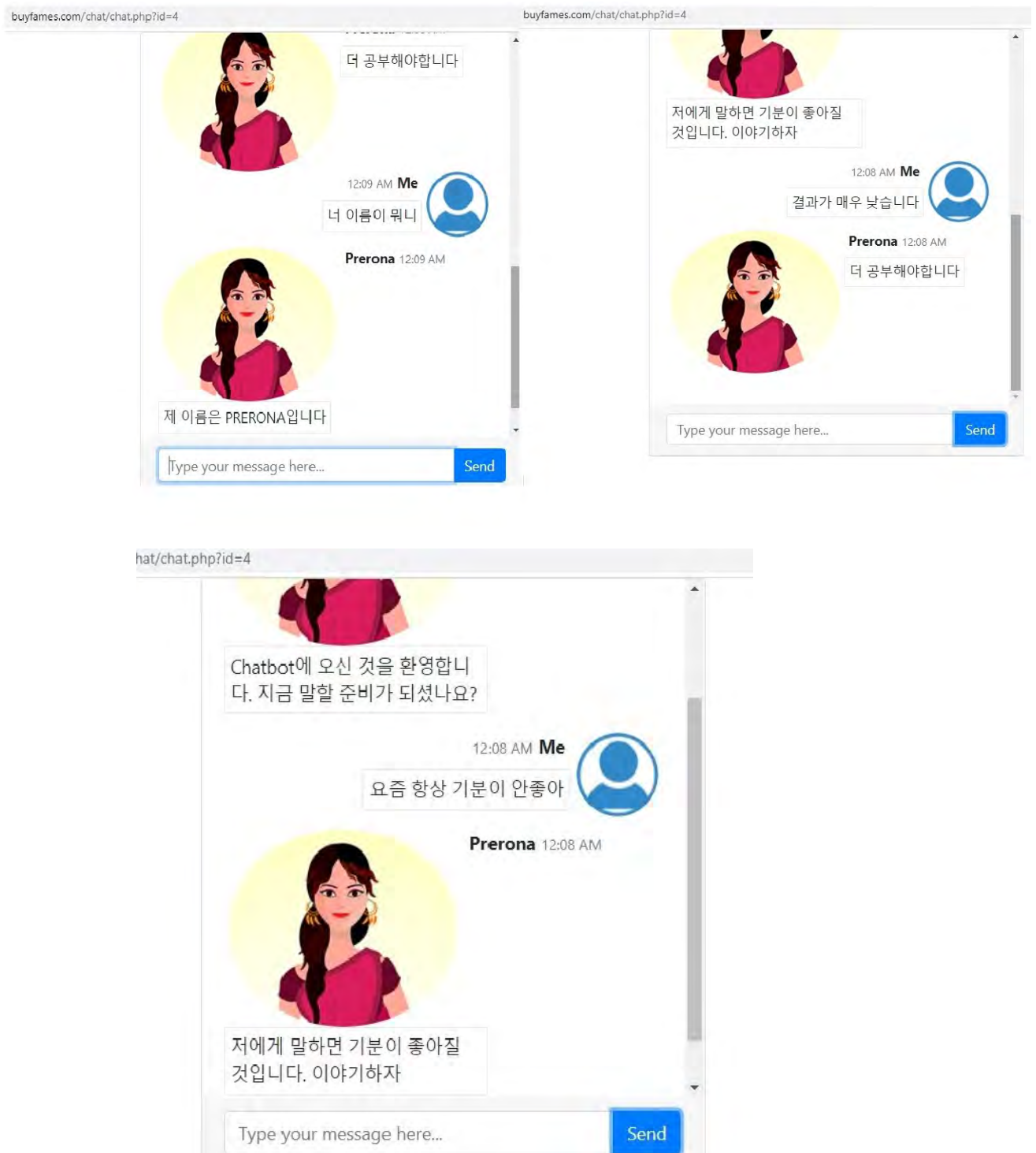


Figure 4.8: Korean User Conversation with Our Bot

Chapter 5

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:

- i. Database of our system
- ii. Typing skills of user
- iii. Data processing efficiency of our system

5.1 Accuracy for Bengali Language

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%. So the other two factors play the main role. 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.

In the bar graph, we drew the result for 30 users in the initial phase, as we mentioned that we tested our accuracy in two phases. In the early phases the accuracy was a little poor. 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 try to collect a lot of Bengali words and those words are added to our dataset. As a result, in the second phase we get some improvements in the accuracy.

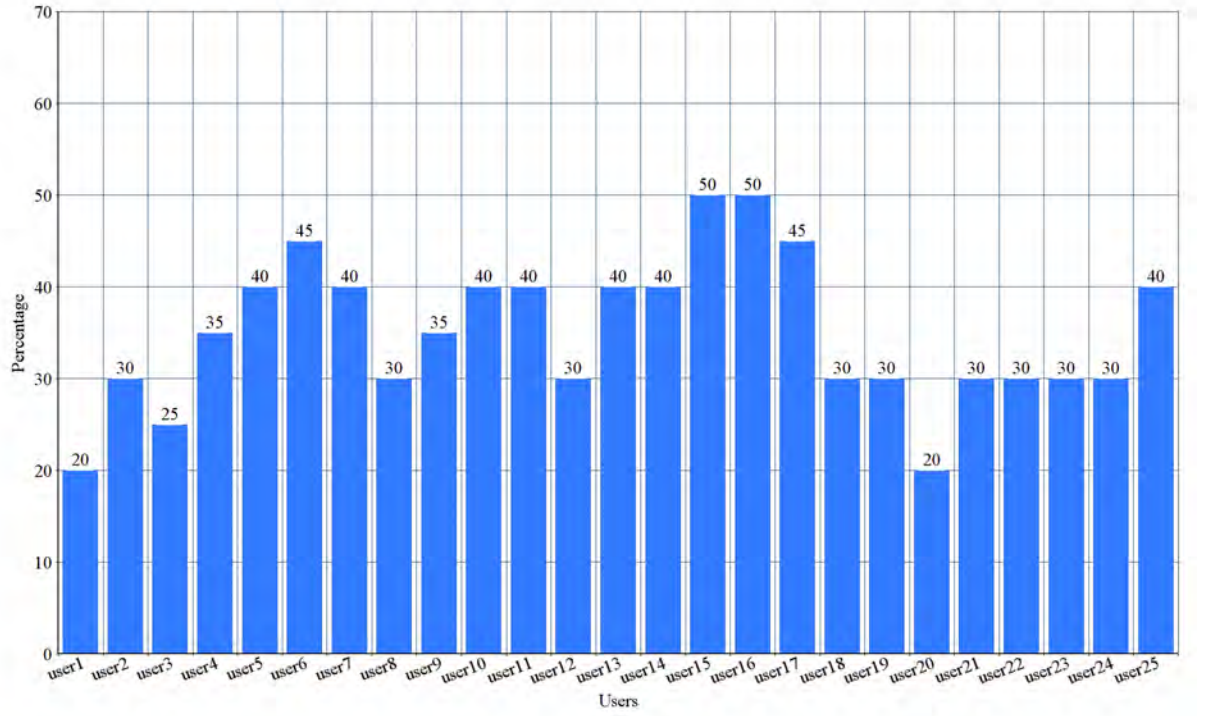


Figure 5.1: Accuracy for Bengali Language (1st Phase)

From the figure 5.2, we can see the improvements in the accuracy. We got highest 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%.

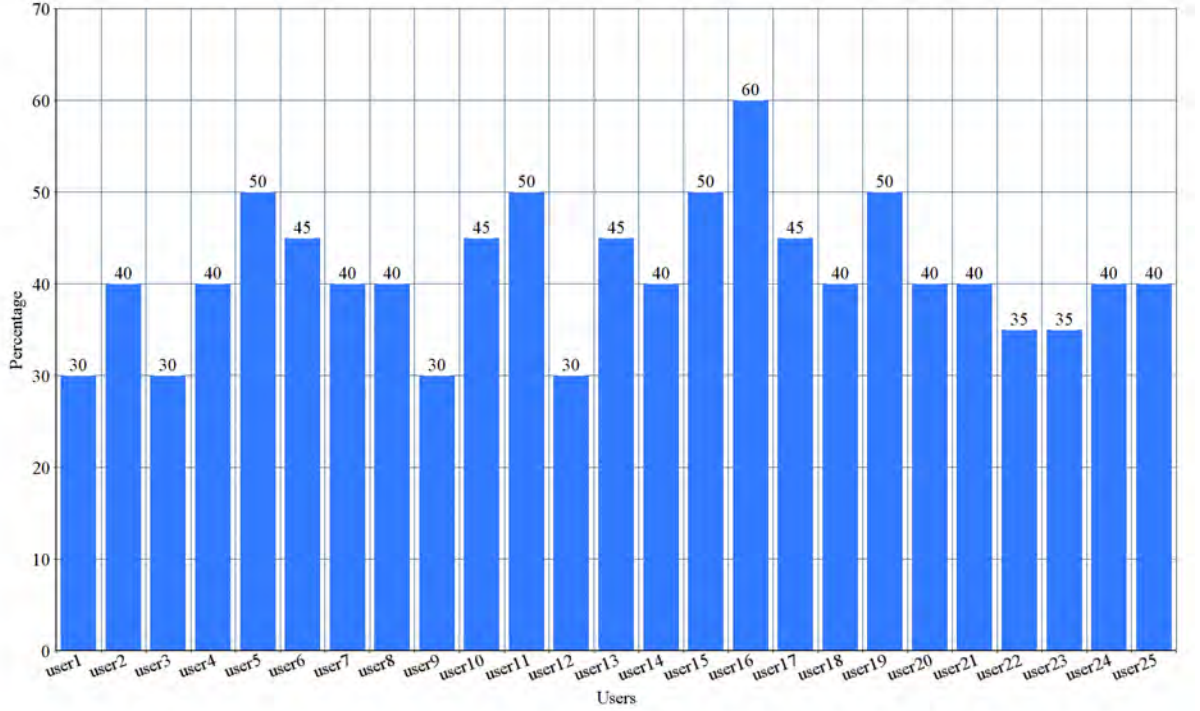


Figure 5.2: Accuracy for Bengali Language (2nd Phase)

5.2 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. He accuracy rate is proportional to the vast-ness of our database. If wrong punctuation and spelling mistakes are reduced from user end, efficiency will increase. 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 5.3, 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.

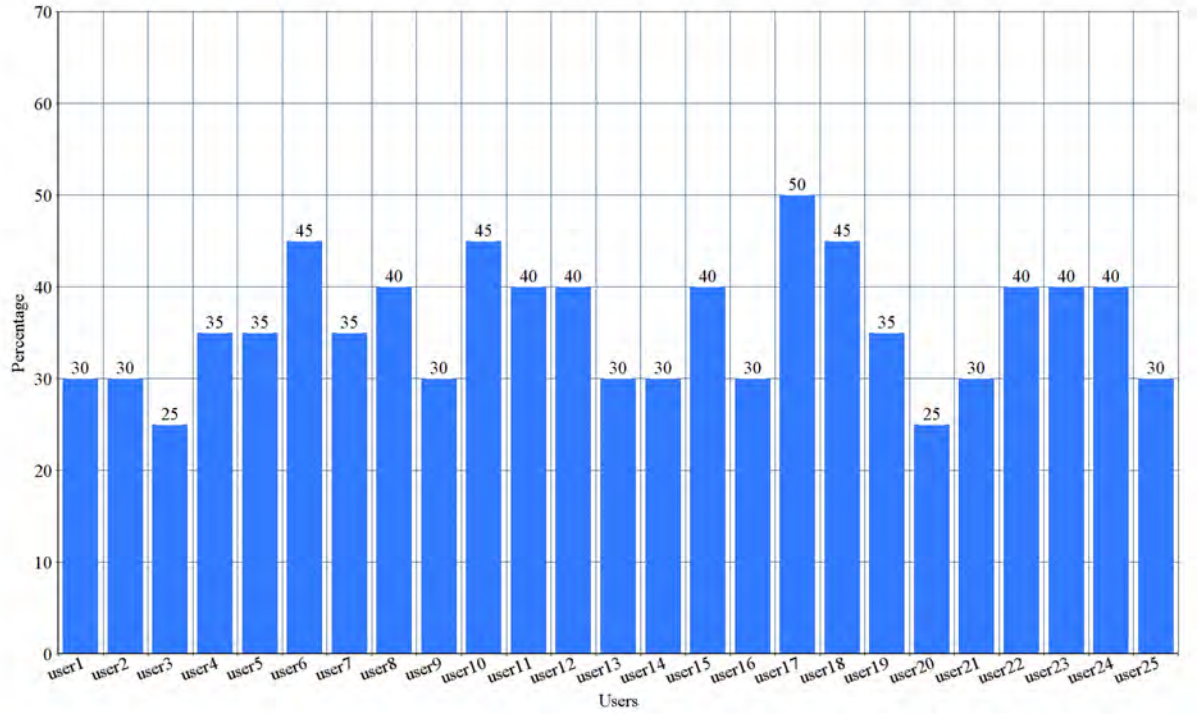


Figure 5.3: Accuracy for English Language (1st Phase)

But in the 2nd phase, we 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 5.4, 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.

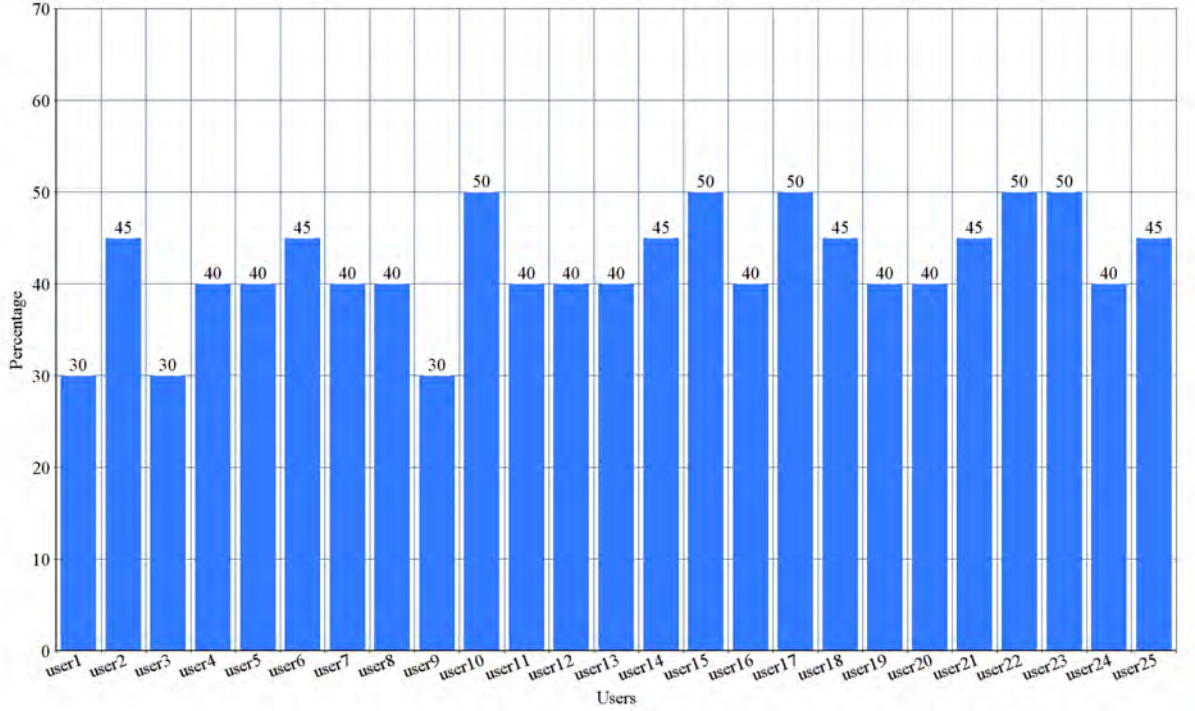


Figure 5.4: Accuracy for English Language (2nd Phase)

5.3 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%. Similar to other languages we try to collect lots of Korean words and those words are added to our dataset. As a result, in the second phase, we get some improvements in accuracy.

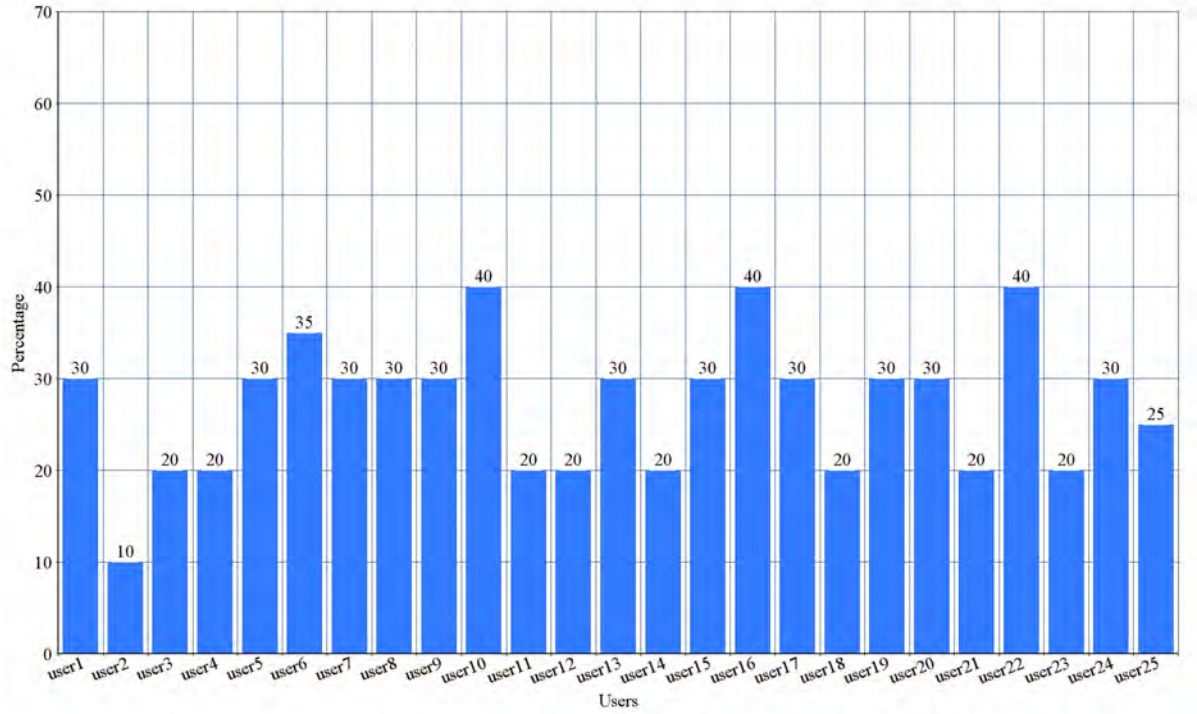


Figure 5.5: Accuracy for Korean Language (1st Phase)

From figure 5.6 we can notice some improvements. From the 2nd phase, we 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%.

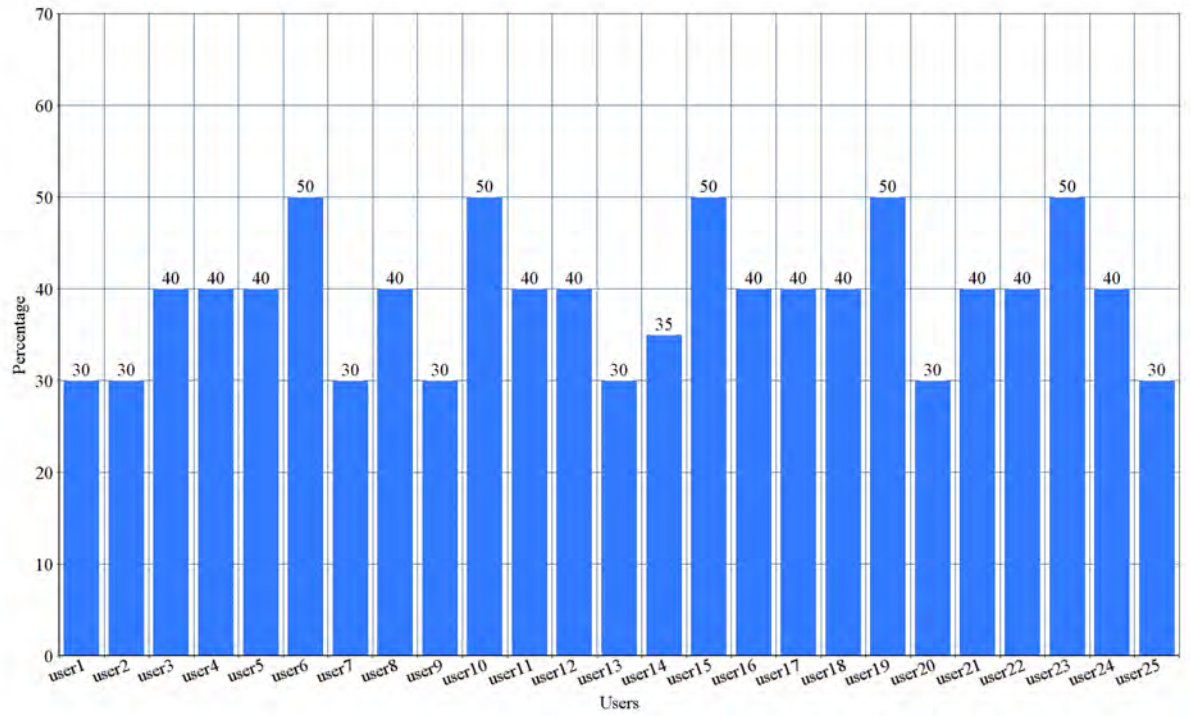


Figure 5.6: Accuracy for Korean Language (2nd Phase)

5.4 Comparison Result Between Phase 1 and Phase 2

Comparison result is shown in figure 5.7

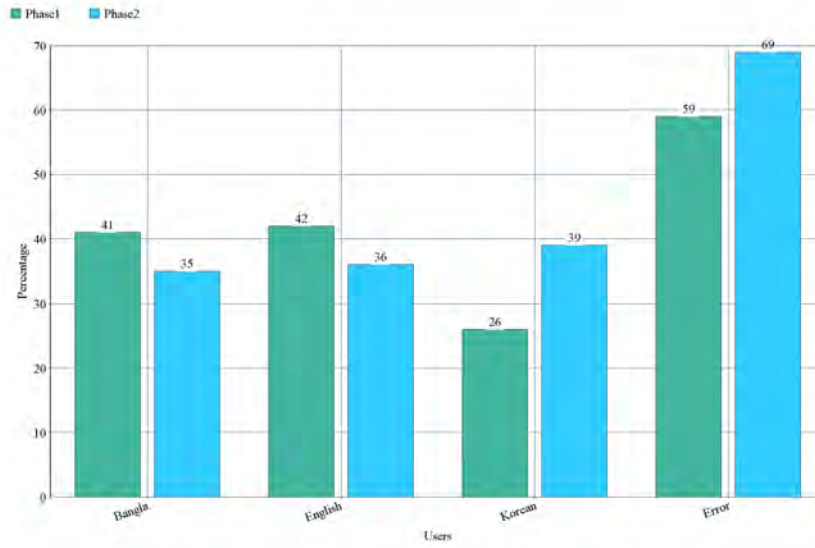


Figure 5.7: Comparison Between Phase 1 and Phase 2

Data Accuracy Percentage for Phase 1

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

Table 5.1: Data Accuracy Percentage for Phase 1

Data Accuracy Percentage for Phase 2

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

Table 5.2: Data Accuracy Percentage for Phase 2

As we can see from both tables, the percentage of error is more in the Phase 1. Our Chatbot PRERONA is an adaptive learning bot so that we can easily update our dataset. We have reduced errors by updating our data corpus.

Chapter 6

Conclusion and Future Work

6.1 Conclusion

Chatbots are widely available and accessible to anyone with a smartphone and Internet access. One of the major targets to build this application is to give mental support to mass Bangladeshi users digitally. Users can get many benefits by using a chatbot for mental health counseling. It can give a user instant information. By incorporating mental health screening tools into a chatbot interface, the user can have a more interactive and user-friendly experience. Research has shown users find chatbots “safe” and easy to talk to [5]. According to the research of Mental Health America – “Depression can occur to anyone, at any age, and to people of any race or ethnic group. Depression is never a ‘normal’ part of life, no matter what your age, gender or health situation” [12]. In Bangladesh, most people are not familiar with the word’s frustration, depression 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 goal of our software system is to impart knowledge about depression to Bengali people through our bots, talk to depressed people and try to suggest use of different sections, in general to help them overcome depression with positive support. PRERONA- our chatbot can communicate with the users/patients through chat and conducts psychiatric counseling. It lets the user express their feelings and acts accordingly. This NLP based chat bot uses PHP, HTML, JSON, JAVA and JavaScript to bring out the outcome. In this research, we present our application software Mental Health Chatbot-PRERONA that can help users and offer consultant contact information through chats. The big challenge of our software system is to create a chatbot that can respond to users accurately. Due to the lack of a Bengali dataset, we have implemented a supervised learning chatbot with adaptive features that will interact with the user based on a pattern-matching algorithm. Our research is a pioneering work in the field of Bengali mental health chatbots. Our Bengali data corpus will help in creating a system for Bengali language processing research.

6.2 Future Work

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 datasets. From our datasets 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.

Bibliography

- [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] 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.
- [4] D. A. Helal, “Depression: Let’s talk”, *The Daily Star*, Apr. 2017. [Online]. Available: <https://www.thedailystar.net/health/depression-lets-talk-1384978>.
- [5] K.-J. Oh, D. Lee, B. Ko, and H.-J. Choi, “A chatbot for psychiatric counseling in mental healthcare service based on emotional dialogue analysis and sentence generation”, in *2017 18th IEEE International Conference on Mobile Data Management (MDM)*, IEEE, 2017, pp. 371–375.
- [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] D. Translations, *Bangladesh*, Apr. 2017. [Online]. Available: <https://www.daytranslations.com/blog/guide/bangladesh/>.
- [9] J. K. Carpenter, L. A. Andrews, S. M. Witcraft, M. B. Powers, J. A. Smits, and S. G. Hofmann, “Cognitive behavioral therapy for anxiety and related disorders: A meta-analysis of randomized placebo-controlled trials”, *Depression and anxiety*, vol. 35, no. 6, pp. 502–514, 2018.
- [10] A. R. Roney, “Suicide on the rise in bangladesh”, *DhakaTribune*, Mar. 2018. [Online]. Available: <https://www.dhakatribune.com/bangladesh/nation/2018/03/27/suicide-rise-bangladesh>.
- [11] 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.
- [12] Mugdhomoyee, “Depression and today’s youngsters”, *Bangladesh Post*, Nov. 2019.

- [13] D. L. Yse, “Your guide to natural language processing (nlp)”, *Medium*. April, vol. 30, p. 2019, 2019.
- [14] M. T. ZEMČÍK, “A brief history of chatbots”, *DEStech Transactions on Computer Science and Engineering*, no. aicae, 2019.
- [15] J. Frankenfield, *Chatbot*, Aug. 2020. [Online]. Available: <https://www.investopedia.com/terms/c/chatbot.asp>.
- [16] *Mental health chatbot*, Sep. 2020. [Online]. Available: <https://woebothealth.com/>.
- [17] *Number of people with depression increases*. [Online]. Available: <https://www.who.int/bangladesh/news/detail/28-02-2017-number-of-people-with-depression-increases>.
- [18] *Your safe space in this difficult time*. [Online]. Available: <https://www.wysa.io/>.