Project and Professionalism

(6CS020)

A1: Project Proposal

<College Chatbot (Heru)>

Student Id   : [2038578]

Student Name   : [Sandesh Thapa – Magar]
Supervisor   : [Susmita Rai]
Cohort/Batch   : 4

Submitted on   : 12/08/2020

Table of Content

Table of Figures

Project Title: College Chatbot (Heru)

1. Introduction

1.1 Academic Questions

I.   How will it solve the college problem to interact with students?

II.   How much accurate is it?

III. Which algorithm is used in your chatbot?

IV.   Can it handle multiple user requests at a time?

1.2 Aims (provide list)

I.   To provide Retrieval-based chatbot for college.

II.   To provide most accurate answer by chatbot.

III. To remove hassle of interacting with every student.

IV.   To provide user friendly chatbot

1.3 Objectives (provide list)

I.   Doing Research on Journal, books, magazines and conferences paper and finding out, how retrieval chatbot is made.

II.   Finding out the similarities and difference by comparison and choosing the best methodology to build it.

III. Making own dataset for the chatbot, to remove the hassle of data cleansing.

IV.   Choosing the right framework and development tools.

1.4 Proposed Artefact

1.4.1 Web Application

There will be a website a for a college to perform the crud operation. Those interested student can submit their details for the admission. Students must be carefully before submitting the details. Since there is no user authentication, crud can be performed before closing the browser. Student can even download the entrance examination paper for the preparation.

1.4.2 Chatbot (Heru)

Main objective is to make chatbot for the college named Heru. It is a chatbot to interact with multiple students at a time. college is facing issue to interact properly with multiple students at a time due to lack of human resources. So, if chatbot is made it can solve those issue. It can even maximize the profit by reducing human resources. Chatbot will be of Retrieval-based to keep it simple. Dataset will be self-made and train, since it is for specifically for a college.

1.5 Introduction of the Project

Organization main goal is to gain profit by providing services and education sector is also one of them. Most of the college and schools have adapted themselves to advertise in a modern way. But they are still lacking in interacting with the user from their website. For people like introvert and who are uncomfortable to talk with people, interacting with chatbot will be better solution. But most of the college don't provided that services. They have to do phone call, but college can't handle hundreds of requests at a time.

At the time of admission, students are in search of better college. They may have seen some advertisement about college, but they want to know more about the college. Hundreds of questions of may be circling their head. So, to clear all those confusion chatbot will be the best solution.

Solving the current problem of the college is main goal as mentioned above. College are facing problems like interaction with individual students. So as a solution, chatbot named "Heru" will be built. In chatbot uses an ask any question related to college. They will be able to get any information without visiting the college from anywhere at any time. If any pandemic happens in future, it will be the best solution to interact with the students. Interested student can even submit their admission form and can book the meeting with college staff for further details. Additional features like entrance examination paper will be available to for the preparation.

2. Literature Review

2.1 Similar finding and Background Research

a. Used of Enhanced Matching Network for Retrieval based Chat bot for multi response.

Insight chatbots are exchange networks that speak with users utilizing regular dialects, where normal human-machine cooperation's are acknowledged with sorts of common language understanding innovations. Neural network-based approaches are the key developments in natural language. Retrieval chatbots are necessary for a variety of candidates to pick the best answer from previous communication statements (Deng, et al., 2019).

	Conversation Context
Utterance 1	A: Hello, I am having problems forcing my laptop to display 1024x768 resolution instead of 1024x600.
Utterance 2	B: What kind of graphic card are you using?
Utterance 3	A: Its intel integrated not too sure of a model number.
Utterance 4	B: Have you tried searching with the hardware drivers' option.
Utterance 5	A: Umm I am sure it returned nothing but I will try again now.
	Candidate responses
Response 1	You are talking about restricted drivers right, right you are.(1)
Response 2	Aptitude is more graphical I think.(0)

Figure 1 Conversation Example [3]

It demonstrates a multi-turn conversation example in the figure one. Since the choosing of the correct answer depends on how the previous topic was spelled out, this example shows that the appropriate answer depends absolutely on the current context [3].

In this journal it is mentioned that to reduce and limit the above problem (EMN) enhanced matching network is proposed. In a matching network, Gated Convolutional Neural Networks (GCNN) are used to create more expressive semantic representation of sentences by removing the Recurrent Neural Network (RNN). In order to collect enhanced interactive knowledge between the answer and any utterance in the background, the enhanced sequence inferences (ESIM) model and local inference modeling are used. Related to SMN, the project had a multi-turn structure rather than a mix of utterances as input. The "Mediocre synthesis" plan would be reinforced in the coordinating organization to permit an essential differentiation among setting and reaction regarding more modest units, (for example, words or expressions). A

consolidated model (EMN-SMN) was proposed to knew the coordinating capacity of the organization. It was then assessed on the Ubuntu Corpus and Douban Discussion Corpus. Also, it was discovered that EMN perform better contrasted and best in class strategies and end up being the best strategy to accomplish the presentation as proposed [3].

Further work was done on response selection, gated convolutional neural network, Local inference and Inference Composition, and Multi-turn structural framework. Early studies have been found to concentrate on single-turn dialog, which only takes into account the last sentence. In the multi-turn answer selection, the whole context was considered where most models combine expressions as a long text to choose the right response. The proposal was made to address the limitation of the recurring neural network inspired by Gated Liner units. Gated Gating Convolutional Neural Network (GCNN) was developed. It was found that (GCNN), with major speed increases, can produce better performance than realistic RNN [3].
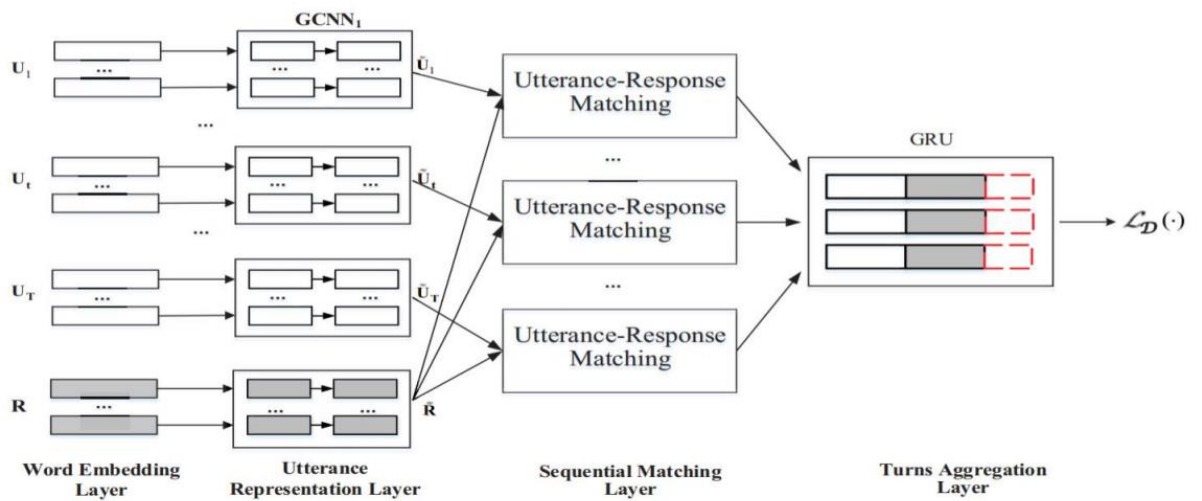

Figure 2 The architecture of EMN-SMN [3]

The EMN-SMN architecture is shown in this figure. The red dashed component displays the corresponding SMN vector. Figure 2 defines the EMN architecture when this portion is removed. A more detailed description of the utterance and response image is shown below in fig 3.
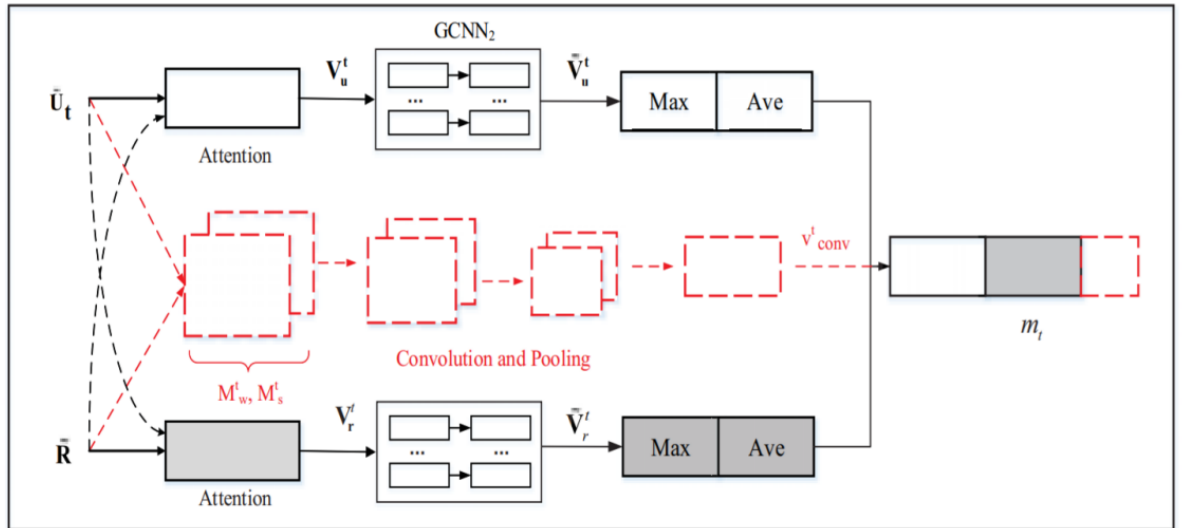
Figure 3 The layer of expression reaction coordinating. The red dabbed line shows the relating layer that SMNN creates [3]

This part portrays the layer of EMN, which comprise of word implanting layer, expression portrayal layer, consecutive coordinating layer and turn conglomeration layer. A query table is utilized in a word installing layer to plan each word to a little vector. GCNN is utilized to encode the info expression and the appropriate response in the expression portrayal layer. Here it figures out how to exhibit for each information word, semantic show. The second-rate creating model with GCNN is utilized in the third layer to build the last coordinating vectors that acquire the expanded data on the distinction between a couple of articulations and reactions. At that point the last journalist vectors will be taken care of into the rotational collection layer where they are aggregated after the successive request of the assertions out of sight in the shrouded conditions of the GRU layer. At last, the comparing score is resolved dependent on the last vectors [3].

As an end creator proposed EMN for multi-turn reaction determination in recovery based chatbots. EMN constructs GCNN's improved semantic portrayal accordingly gathering further reliance information through nearby surmising displaying and arrangement of derivation. Notwithstanding that the creator have proposed to utilize EMN-SMN consolidate to accomplish improved effectiveness by incorporating the qualities of both EMN and SMN.

b. Stacked Multi0head Attention for Multi-turn Response Selection in Retrieval-based Chatbots.

In the field of deep learning, response selection is the key to retrieval based chatbots. It is necessary to capture semantic dependencies in different segments in utterances for the retrieval chatbots. It has been found in early studies that selection of responses only takes account of the latest argument, which is called single-turn reaction selection. But, looking out our day to day conversation, it's been important to extend single-turn response selection (Yu, et al., 2019).

In the further studies it was found that attention mechanism as an essential part for sequence modelling. The sequence modeling is unique in its ability to model a phrase without taking into account the location of every word in the sentence and its advantages over encoder-decoder [6].
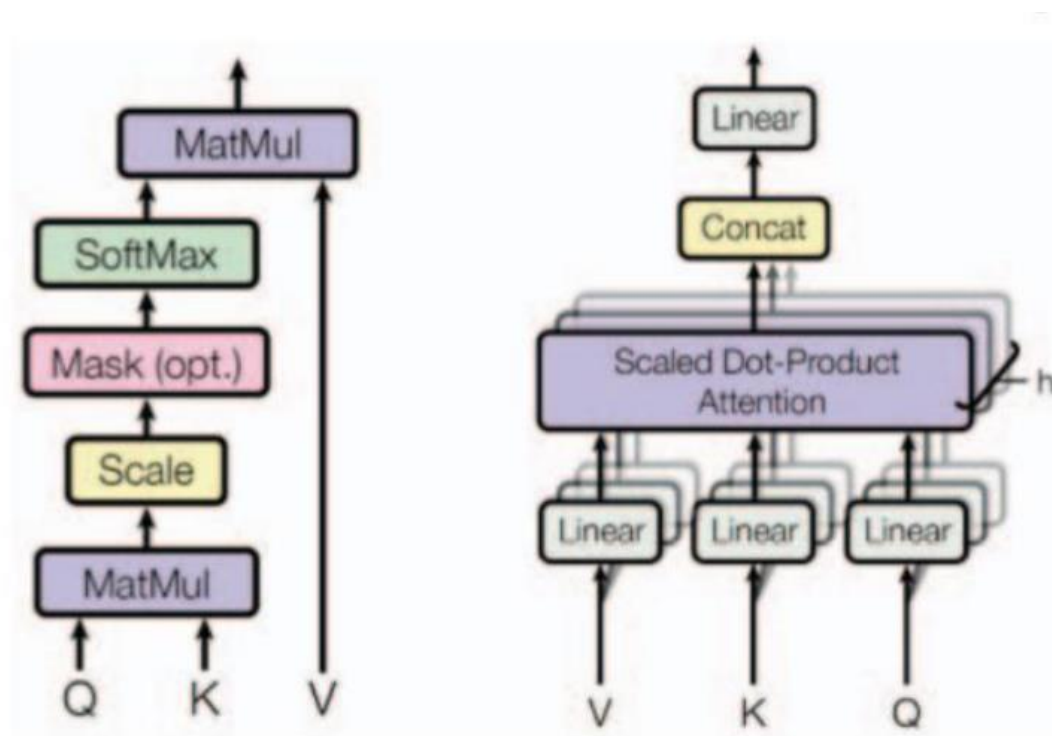

Figure 4 Scalded dot-product attention (left and multi-head attention (right)) [6]

This figure 4 shows the attention mechanism used in Transformer include dot product attention and multi-head attention. We can see in the figure input consist of three vectors, where Q is denoted as query vector, K as key and V as value vector where K is equal to V. Formula is defined as:
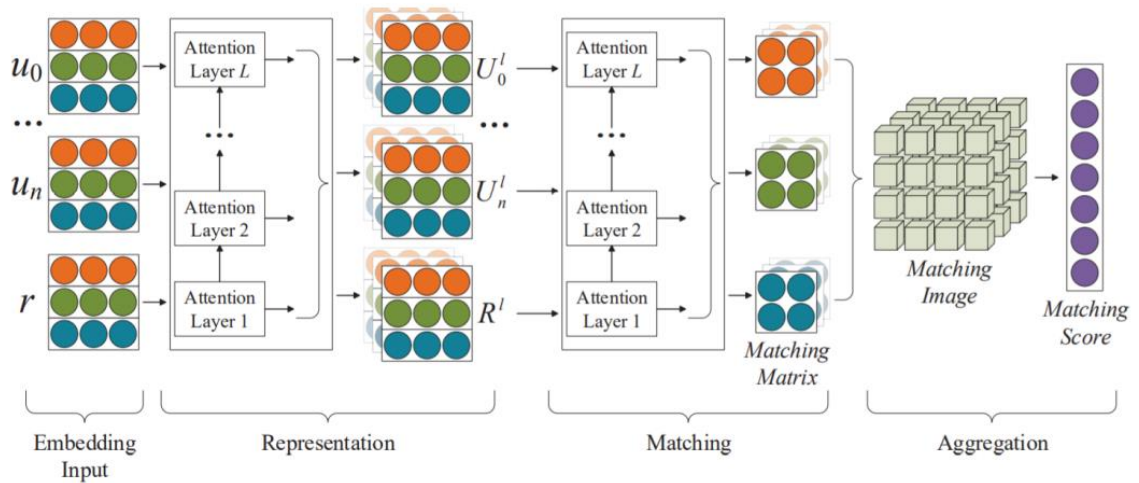
This figure consists of three units: representation unit, matching unit and aggregation unit. The term embedding tensor for input sentences is designed to produce multi-turn representation, both for a multi-turn context and for candidates' responses. The matching unit integrates the candidate's response to multiple-turn context representations via another multi-turning attention stack, and achieves matching matrices with the multi-turning framework The aggregating unit conducts convergence and max pooling operations with matching matrices, collects more secret knowledge about matching features between the multi-turn context and the candidate answer and then integrates these in one perception layer, meaning the association between each context and their candidate response [6].

To conclude, the author has suggested chatbots to stack multi-head attention with a multi-turn context. Based on DAM, the author extends his attention mechanism and only then can they get more information in multiple-head in various representation areas, substitute the single-headed attention from the transformer. Thanks to this ability to capture the long-distance dependency of sequences is to some degree enhanced.

c. Knowledge Based Chatbot with context Recognition

Chatbots or automated replying bot whatever it is named is used to interact with the humans through Natural language. With the increase in popularity of chatbot, we can see it in places like customer service, collecting information and other dialogue systems. Bot uses the NLP techniques to understand our natural language where some bots use string manipulation techniques to response. (Wijaya, et al., 2019)

In this journal article, writer is mentioning to develop knowledge based chatbot with context Recognition. Writer main aim to develop this bot was to response the answer that were frequently asked by the new student at the time of admission. For the development of chatbot following research were carried out, like text mining method to be used to improve the ability to handle the word synonyms and stemmed stopword dictionary. In the process of text mining additional results were produced which acts as a synonymous of dictionary for learning. Upon processing of the information, the basic knowledge of chatbot will then be used to perform the best response calculation by means of context recognition to filter relevant information. Then the similarity of binary cosine is determined and the best results are determined as an answer [7].
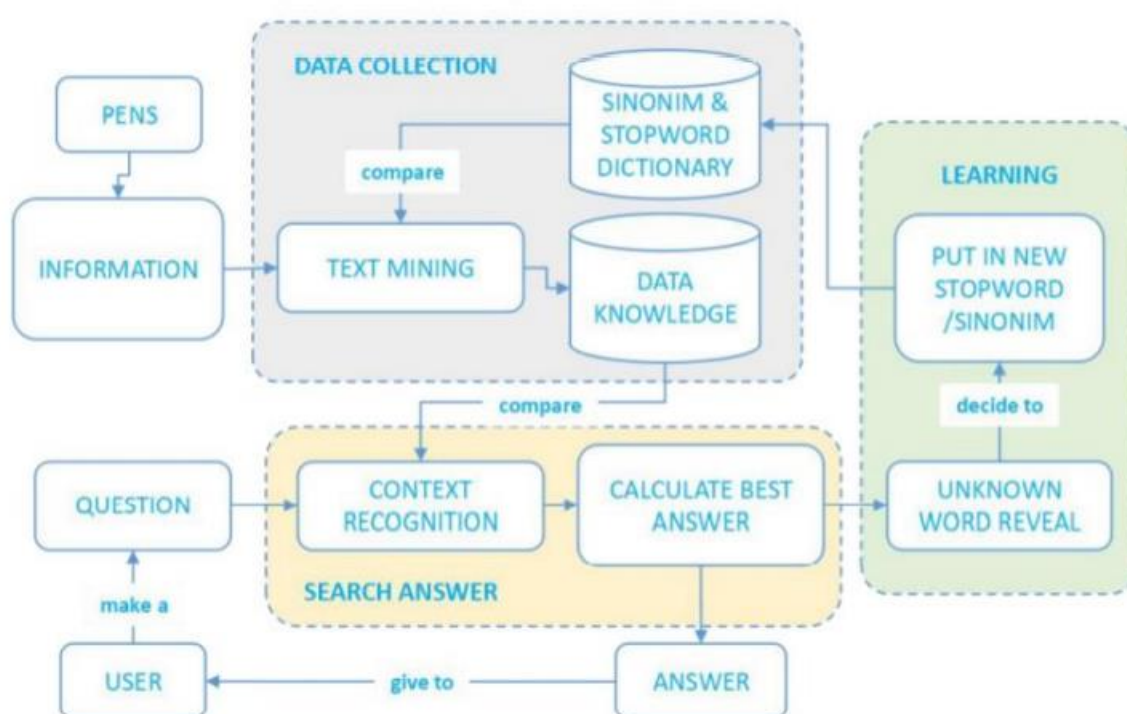

Figure 6 Design System [7]

In the above figure 6, it is the proposed design system of the chatbot by the writer. It is mentioned that it will be used in the webpage of pens.ac.id and calling through the JavaScript.

After data collection was done from PMB (Penerimaan Mahasiswas Baru) PENS, then data were processed with text mining to get parts that can be computed by the system. In text mining method following function were carried out like casefolding to equate and deal with uppercase and lowercase letters for eg: Before : "Apa kabar Warga PENS?" After : "apa kabar warga pens". Tokenizing to break the sentences

so that it increased to understanding words, phrases and narrative paragraphs like:

Before   : "Apa kabar Warga PENS?"
After   : "apa, kabar, warga, pens,?"

To increase the accuracy regex remover is used to remove the symbols because they are useless in the sentences like question marks, exclamation marks, symbols, points and others. For eg: dash is removed as show in this figure. Stemming is being used to change the word to a basic word because often in a sentence there are affixed words. Similarly, other functionality like spell correction and stopword is also used [7].
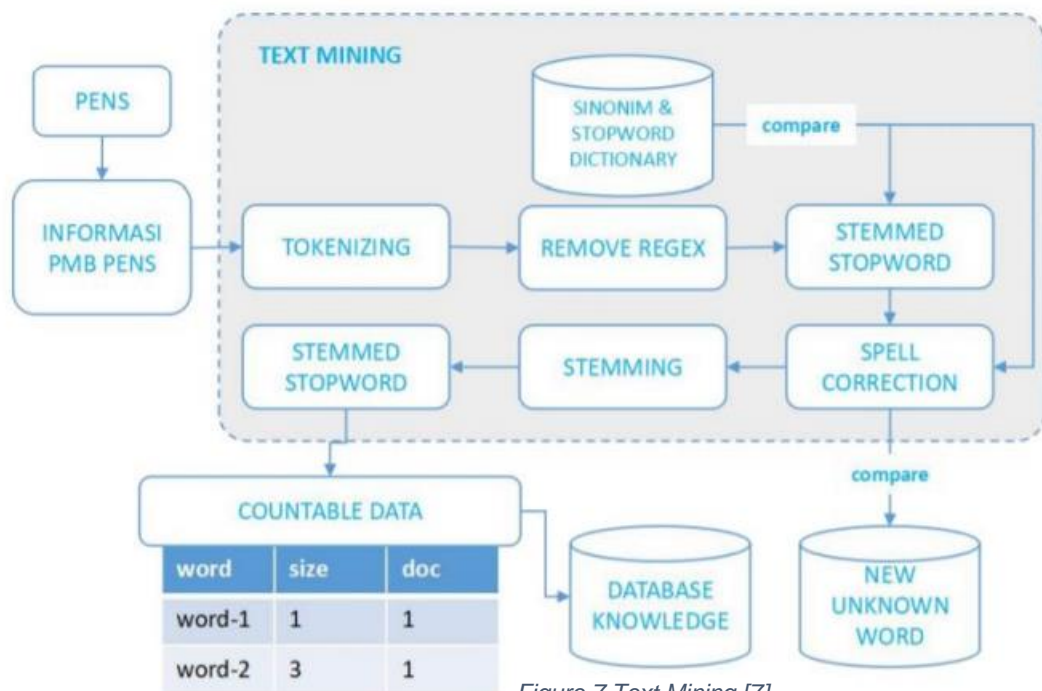

Figure 7 Text Mining [7]

In this figure 7, it shows the above-mentioned functionality of a text mining. It shows how the text mining works. For future use the result of text mining will be entered into the database. In the process of text mining new foreign words can be found and will be entered to the database.
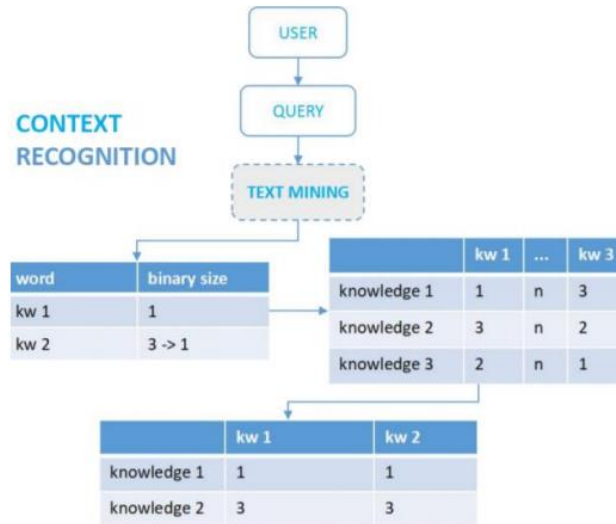
Figure 8 Context Recognition [7]

Since the training data that is entered will be in the raw form, it needs to be preprocessed first. And these processes should be done before processing and text mining. It can be done through the stopword removal, tokenizing and others. The only advantage on chatbot that writer or developer can perform is by improving the speed of a chatbot. It can be done with context recognition; it takes only the required data [7].

To differentiate the two sentences like which one is the answer and other question, binary cosine similarity is being used. It is a way to quantify a vector discrepancy between two current data as show in the formula:

$$similiarity = \cos(\theta) = \frac{A.B}{\| A \| \| B \|} = \frac{\sum_{i=1}^{n} A_i B_i}{\sqrt{\sum_{i=1}^{n} A_i^2} \sqrt{\sum_{i=1}^{n} B_i^2}} ...(1)$$

Explanation :
A : First dataset
B : Second dataset
i : index

To conclude, the author has changed or modified the preprocessing style according to the need and to find out the appropriate answers for the chatbots. The author also demonstrates how effective the usage of dictionary-based synonyms is and also how the input is manually used. The author shows in the pre-processing that the use of synonyms manually improves the precision and expertise of bots, but it requires a long time and deep comprehension from the admin to construct dictionary synonyms. The authors compare the responses received using six measures with each distinct approach in the process of seeking answers. The effect is that dictionary synonyms can be used to improve precision by up to 25% (Wijaya, et al., 2019). But for chatbots, binary

systems for cosine similarity are strongly recommended because they have a clear objective from user requests on chat forums (directly mention the root of the issue) because how often the terms are not for computation, but rather emphasis on the collection of keywords in the issue.

d. Multinational naive Bayes Algorithm with Logistic Regression to classify the intent in chatbot

It is not physically possible for a human to reply all the conversation. With modernization of the technology people are shifted to modern solution to deal with the conversation problem. For the growth of any business, continuous interaction is needed, so a solution chatbot will be the best to deal with it. Big company like Facebook, amazon, Microsoft provide their own chatbot platform but there is limitation where the developer can't customize as they like. So, building own chatbot from the scratch is the better way.

In chatbot, user-bot interaction is accomplished by natural language, such as: computer conversation systems, robotic assistants, dialog systems, and chatterbot. In this journal author have proposed to make a chatbot to simulate a human conversation. To identity the intent author has proposed a classification method instead of user input or socalled intention classification. As a classification method Naïve bayes method is used and compared with logistic Regression method to find out the class intention (Helmi Setyawan, et al., 2018).

In related work author have gone deep in several text classification and chosen Naïve Bayes among them. Naïve bayes is a system of classification based on a model and provides competition for comparing text categorization classification results. Due to simple and highly efficient Naïve Bayes classification algorithm is highly used. The classification is used as Bayesian the learning method of probabilities and each categorized algorithm function is not dependent on other features values. The algorithm of Naïve Bayes can be adaptive and intelligent, functional but also satisfies individual requirements and is therefore widely or extensively used in commercial application [11].

General liner relapse or calculated relapse whatever it's named, permits us to test the impact of mathematical components on double reaction. Between subordinate variable like (Y) and autonomous factors (X0,X1,X2,… … .Xn), relapse center around their relationship. 'Y' is a high level an incentive as liner relapse happens when coordination's have a discrete worth. In the calculation

of the strategic models, which have any an incentive from negative to positive limitlessness in a limited measure of info and yield, the calculated condition is otherwise called sigmoid capacity. Since relapse can be utilized in AI applications, it will grasp vector esteems and compute the info variable and coefficients and gauge the class spoke to as a word vector esteem [11].
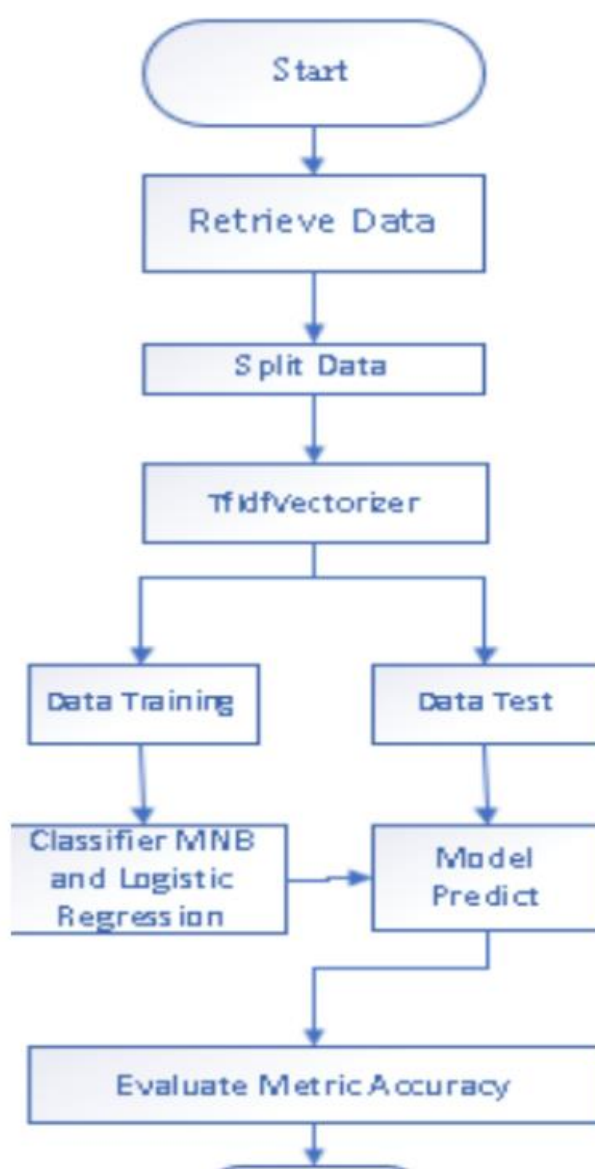

Figure 9 Evaluate Model [11]

In this figure it decides the Intent arrangement of the chatbot framework to be constructed.

From this study we can find that researchers continue to use training data in this analysis with a relatively small volume. It is possible for errors to occur when preaching fewer training data to the class intentions in any class of intentions. The specificity of the two strategies can be differentiated when the training data are used and the sum of data in each intention class is the same. It was additionally discovered

that when the purpose class had a similar measure of information in every unit, there was decline in the precision of the calculated relapse models [11].

e. Web Based UNIBOT for University

In this journal chatbot named (UNIBOT) will be made. It is almost similar to system that was proposed to developed in FYP. The developers are trying to reduce the hazard that arises at the university by developing the UNIBOT using their own algorithm, it was mentioned that algorithm like ALICE bot which uses AIML and program Eliza will be taken as a reference which uses pattern to be matched (P.Patel, et al., 2019).

From this journal we will get to know that developer is making their own algorithm. Making own algorithm will be time consuming and not sure either it will work or not. Therefore, already built in algorithm will be better solution to build the chatbot.

f. A Comparative Analysis of Algorithms for a School Based Chatbot Using Natural language Processing in Textual Analysis

Here the team of developer is developing a chatbot named Melbot to compare the widely used Algorithm (NLP) Semantics and Tokenization Analysis. Summing up the Machine learning and Natural language processing algorithms the developers are trying to find out how textual data is processed at the time of passing through it and its nature. In machine Learning developer is using regression algorithm. It is used for forecasting and discovering links between variables out cause and effect. Developers are even developing different chatbot to know which one will the best for Eg: chatbot using Semantic, tokenization and mixed (Jeremie, et al., 2018).

From this article, we will get to know about the NLP and machine learning in depth and which algorithm will be faster or combining both algorithms will be more accurate at text processing.

c. Chatbot for college, with A.L.I.C.E

In this Journal, developers have mentioned about creating a chatbot using ALICE (Artificial Linguistic Internet Computer Entity) to enquiry about the college. Main purpose to build this project was to guide the new students who are facing problems at the time of admission. Developers choose to build their

own chatbot using the ALICE and AIML which is freely available under GNU General Public License since buying the chatbot was expensive (Bani & Singh, 2017).

d. Intelligent Chat Bot for Banking System

In this Journal article, developer is mentioning about making Highly Intelligent chatbot for banking system. This Intelligent chatbot will answer almost all answer and even give suggestion while taking the input from the users. It can even take input in the medium of voice or written format. Since, it will be developed with highly trained AI, there won't be any language barrier for chatbot to understand the human language. Developer will be developing on C# with .net framework and for converting voice to text, speech recognition framework   and for text to speech, Speech Synthesizer will be used (Dole, et al., 2016).

From this journal article, we can get to know how the input is taken in the form of voice from the users. If this feature is added then it would be better for the users but it is complicated to implement. So, a simple chatbot with input text will be better.

e. Chatbot using TensorFlow for small Businesses

In this journal article, writer is mentioning about making chatbot for the small business to grow up. He has potentially seen education sector as one for the good area to implement of a chatbot (Singh, et al., 2018). It was found that in the journal most of the chatbot related to college or school was made for the inquiry, which is similar to my system. Neural network plays vital role in the chatbot as a model so he planning to use TensorFlow and NPL to maintain the context of the conversation [**Error! Bookmark not defined.**].

2.2 Analysis of the similar System and Comparison

Note: EMN is known as (rep&match-GRU) and GRU as GCNN

In enhanced matching network, further investigation was done to know the effects of different parts of EMN. When comparison was done between GCNN and GRU in the utterance representation layer it was found that performance of EMN (rep&match-GRU) was decreased compared with EMN. So, we can say that in catching the semantic information GCNN outperformed GRU. Moreover, it was seen that, marginally decrease in the performance of

EMN when GCNN was used to learn the sequential order of sentences in aggregation layer. Even if EMN (rep&match-GRU) efficiency decrease in relation to this compared to DAM which is similarly based on a multi-turn structure, EMN still has marginally improved. Even though they have their pro and cons, to learn chronological information of utterance GRU is considered better while for semantic representation GCNN is consider more suitable. Overall, it shows the significance of local inference and inference composition in EMN.

In stacked multi-head attention retrieval chatbot, when experiment was done in ubuntu corpus and Douban corpus, it was seen that the performance of their model was improved in most metrics. And since the ratio is undefined between positive and negative in the Douban dataset which will result in low score so, MAP and MRR for comparison are suggested. In overall DAM outperformed by the margin of .96% in terms of R10@1 on Ubuntu Corpus v1, 1.0% in MAP and 0.3% in MRR on the Douban Corpus.

To do comparison between multinomial naïve Bayes algorithm and Logistic Regression for intent classification the author made the chatbot. After the evaluation was done in both naïve and logistic for the intent classification, it was shown that the accuracy of naïve is 0.6363636363636364 and of logistic 0.0909090909090909. Form this data we can clearly see that the logistic regression is more accurate by 12.5%. Also, at the time of evaluation precision and recall were obtained and it shows that logistic took average of 0.59 for precision and 0.73 for the recall, where bayes took 0.53 and 0.64 respectively to produce the data. In overall, using the logistic regression method will provide more precise and higher precision value compared to Naïve Bayes. Therefore, it is safe to conclude that the logistic regression efficiency is higher than the Naïve Bayes model

In this University chatbot (UNIBOT), ever thing related to chatbot were explained and presented in detailed. From details design to the algorithm that is begin used and about implementation. After the completion of chatbot, the student doesn't need to go college to gather the information. They can access the information from anywhere at any time.

In College Enquiry chatbot ALICE was used to made it. ALICE used a basic pattern design in contrast to Elizabeth, representing input and output of pattern matching algorithms, of Elizabeth using the input rules, patterns and

output rules for responding. The recursive approach is utilized in Alice to simplify feedback by recurringly naming match categories.

In chatbot using TensorFlow, when method was demonstrated, to some extent it was successful for small domain size . But with increase in size of the intent file it will be more accurate. So, we can say that the precision of the bot corresponds directly to the size of the intent file used for the formation of the bot.

3. Project Methodology

Current problem of a college is that they are not able to interact with their students properly. Those students who are already enrolled in the college don't get enough information from their college. They have to either depend on their college mail or phone call for major updates from the college. If chatbot is made then, students can ask information regarding their studies and get the helpful link from the bot. For the new students who want to know about the college in detail is not physically available now. Answering the phone calls and making them understand fully is not possible, so chatbot will be the better solution. It will help to reduce the human manpower and cost of the college. Accuracy or accurate of a chatbot response depends on how the chatbot is trained, the more you gave information in the dataset of a chatbot the more it will be accurate depending on the which algorithm is used. Proposed chatbot will be made by using python library chatterbot which includes Search algorithm, naïve Bayesian classification algorithm and NLP to understand the natural language.

Since proposed project is not that big and does not require users to login or sign up so I am using Django or Flask as light a weight framework with python, and chatterbot python library will be used to build it. It will include all the requirements to build the chatbot, for Eg: it will get the input from the sources like console, api, speech recognition, text. Then it will process the input by the multiple logic adapters by matching the closely matched input statement and return a response that is closely matched with highest confidence on the matching to the input statement (Gunther Cox Revision, 2019).
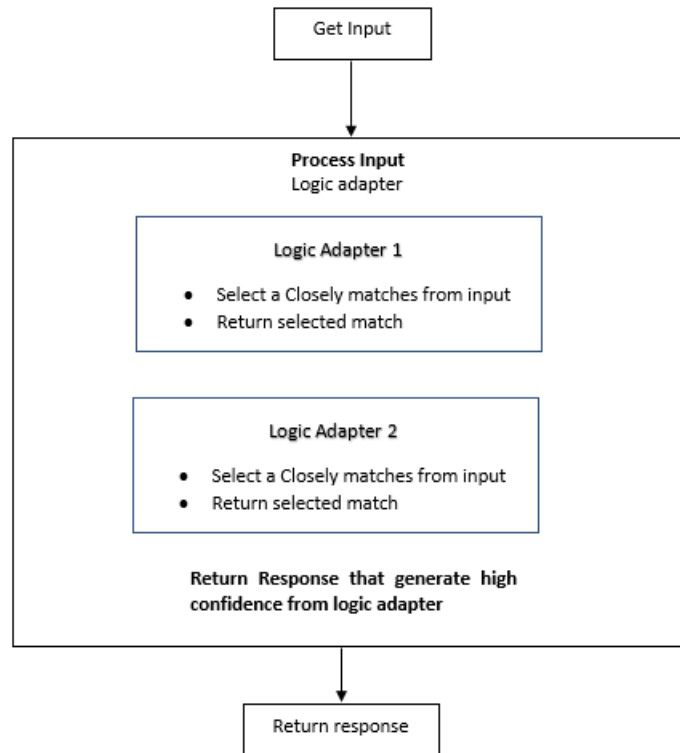
Figure 10 Working flow of the chatbot

3.1 Methodology

The incremental construction model is used for this analysis. In this methodology, research will be done in the parts and built it which will then be interconnected. It is a software development approach in which the product is partly developed that means elements need to be developed and maintained to ensure that connections at the testing process are successful. Scrum Production process is the approach used throughout the methodology. Since this is the FYP all the features will be done in individual, supervisor will act as a scrum master. The key method to scrum is to use gradual sprints. Any sprint will incorporate the features or tasks specified for the project and should be done by the end of each sprint. A final Gantt charts will display the outcomes of each sprint and other work like developing backlogs, planning sprints, using a scrum board, doing meetings with scrum master(supervisor) and sprint review will be done.
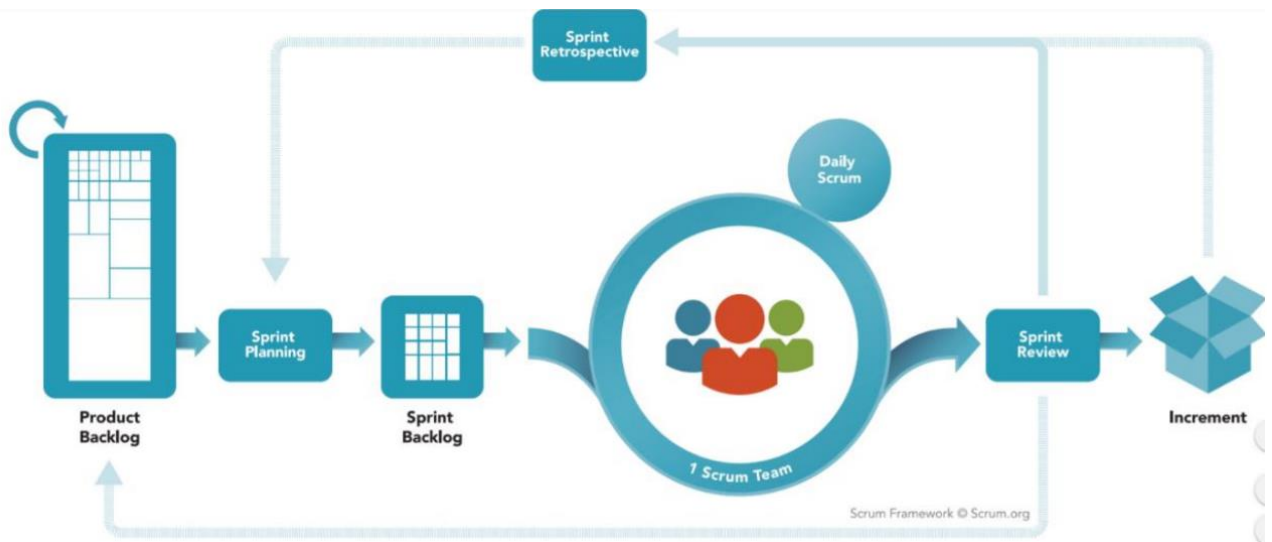
Figure 11 Scrum Development Method

I. **Backlogs**

It is the list of functional requirements that needs to be address at the time of development. It may change time to time with the users need.

II. **Sprint**

In this phase, the work is carried out to achieve the requirement of the users according to their functional requirements. It may last up to one month usually.

III. **Scrum Meeting**

This is the meeting between the scrum master and the team members to focus on answering the three major question "What did you do since the last team meeting?", "What obstacles are you encountering?" and "What do you plan to accomplish by the next meeting".

IV.   Demos

After the certain portion of software is developed, to take the evaluation form customer software is delivered according to their functional requirements

4. Tools and Technique

Hardware resources:

I.   Desktop/Laptop

Both desktop or laptop is fine for the development. For portable latop if preferred. Specification is minimum, RAM 4GB, hard disk space 20 GB is enough. NO GPU required, CPU is fine and processor above i3 with 5th gen is suitable.

II.   Internet Connection

Internet connection is needed at development time only to import the python libraries for the chatbot. And if the website is hosted online then it requires internet connection. It does not require high Bandwidth; it work on normal ADSL internet connection.

Resources

I.   Python

Python will be use as a programming language. It has many built in libraries. It will be easy for the developer to build the web or app development. It is easy to use comparing to java and the for my project it has pre built in libraries as an advantage.

II.   Django or Flask

Since my project is not that big, so I am using Django or Flask as a framework according to need. It is built in python so it will be good enough for CRUD operation.

III.   GitHub

I will be using both GitHub as a software for the version control and for the backup.

IV.   Sqlite

As a database, Sqlite will be used. It is used to store the train data set. Data set may be in the format of json or normal text file.

V.   Html/CSS

Since it is web based chatbot, basic html with CSS will be needed for the frontend. And little JS for the backend.
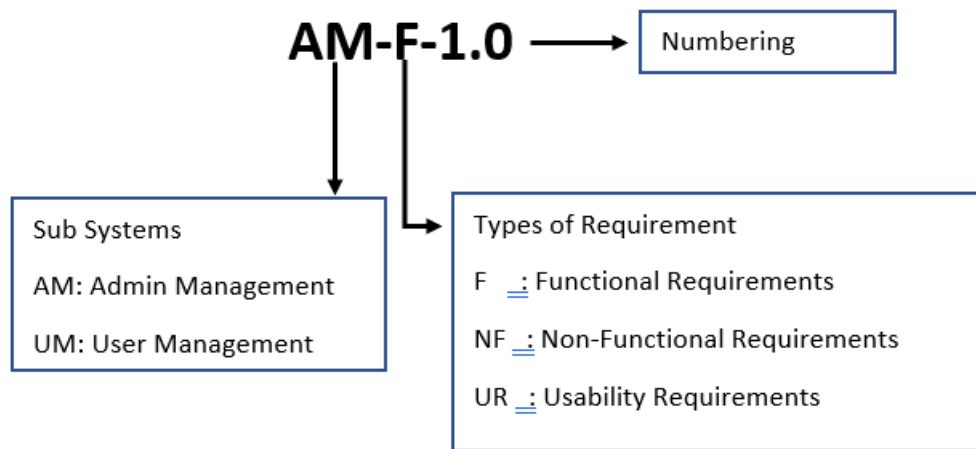
VI.   Vs Code

As an editor Vs code will be used for coding. It is lite version of Microsoft, which support almost any type of language. VS store have many extensions which will be needed in software development.

VII. Draw.io

This online software will be used as SAAS. Here all king of diagram will be made like ERD, sequence and others.

5. Artefact Design

5.1 SRS

Req Code	Req Desc
AM-F-1.0	Admin also should create the account of every users.
AM-NF-1.1	Admin should use unique username for each user while creating the accounts and password should be hashed while storing it.
AM-NF-1.2	The email address should be unique to each user.
AM-NF-1.3	The password length should be equal or greater than 8.
AM-F-2.0	The system should allow the admin to delete users and make some update in users details as per required.
AM-F-3.0	The system should allow admin to view the user's details.
UM-F-1.0	User should be able to sign up on their own.
UM-NF-1.1	New sign-up users should be on hold, until approve by the admin.
UM-U-1.1	When new user sin-up they should be redirected to hold on page, with certain message.
UM-F-2.0	Users should be allowed to chat with the chatbot.
UM-NF-2.1	Non registered users also should be allowed to chat with chatbot.
UM-F-3.0	The system should allow user to login.
UM-NF-3.1	Session should be created after each logged for the individual users.
UM-U-3.1	User should be redirected to dashboard page after the successful logged in.
UM-F-4.0	System should allow users to download the entrance paper.
UM-NF-4.1	Only the registered should be allowed to download the entrance papers.
UM-F-5.0	System should allow valid email users to send feedback.
UM-NF-5.1	All users with valid email should be allowed to send the feedback.
UM-U-5.1	Message should be shown on successful submission of the feedback.

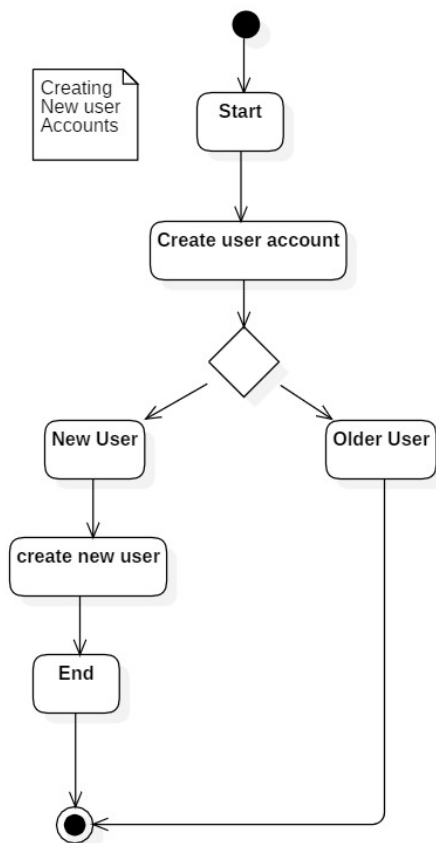## 5.2 All Designs and Related Diagrams

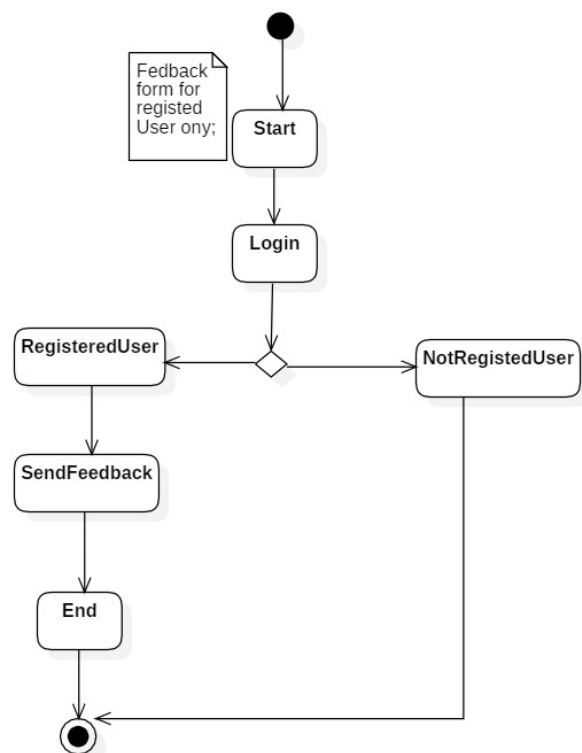### Activity Diagram

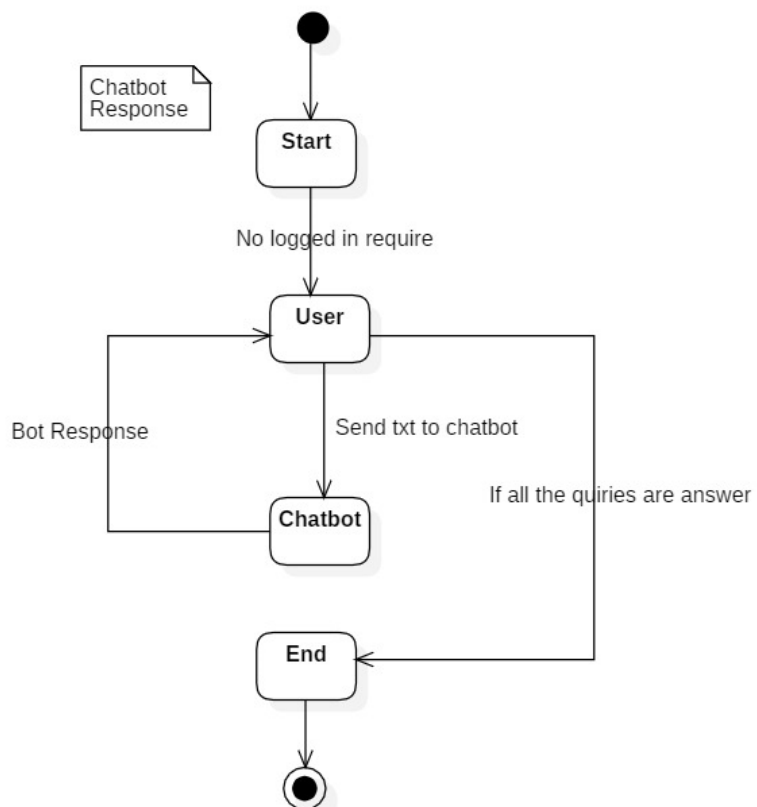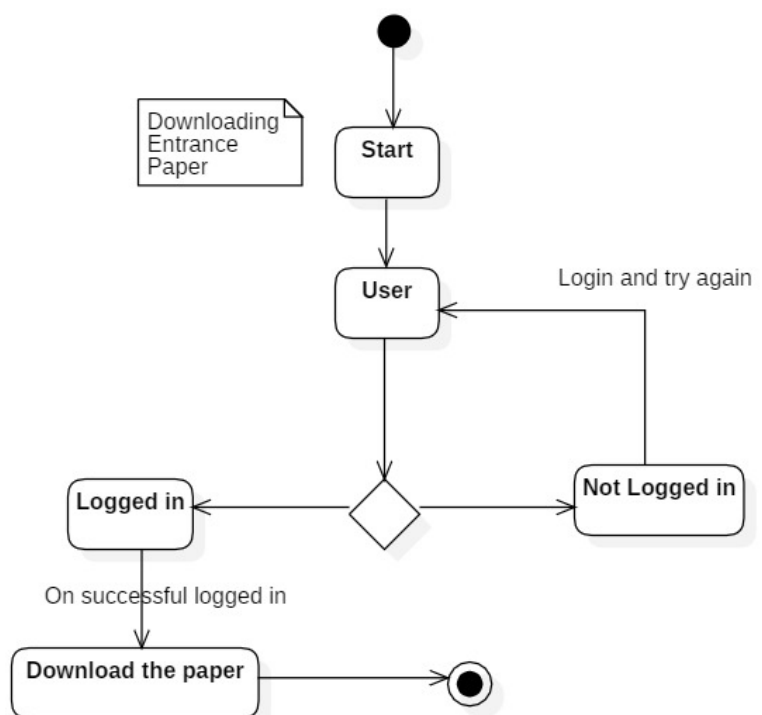*Figure 12 CreateNewUser*

Figure 13 Feedback form for user

Chatbot
Response

Start

No logged in require

User

Bot Response

Send txt to chatbot

If all the quiries are answer

Chatbot

End

Figure 14 Chatbot

Downloading
Entrance
Paper

Start

User

Login and try again

Logged in

Not Logged in

On successful logged in

Download the paper

Figure 15 Downloading Entrance Paper
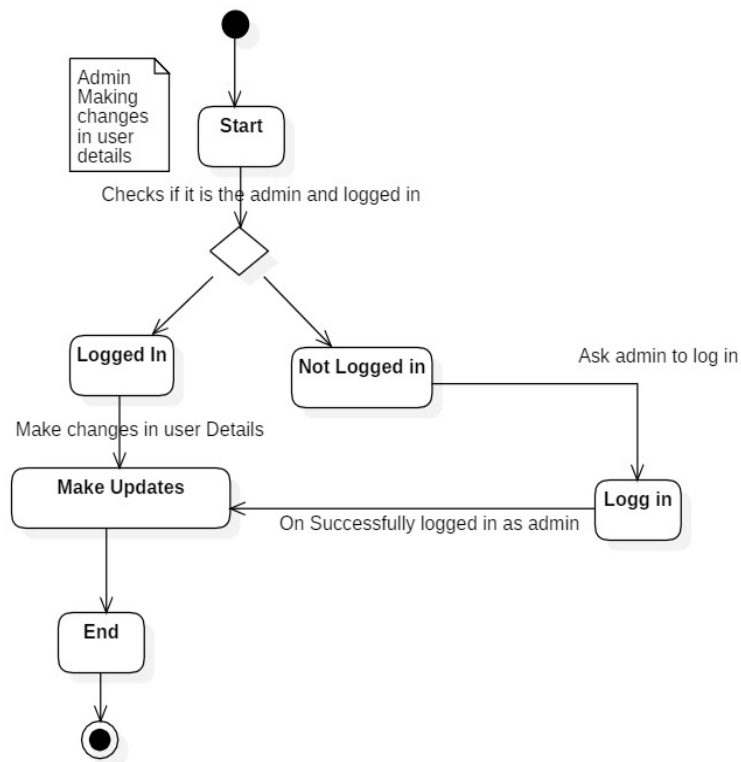
Figure 17 CRUD for Admin
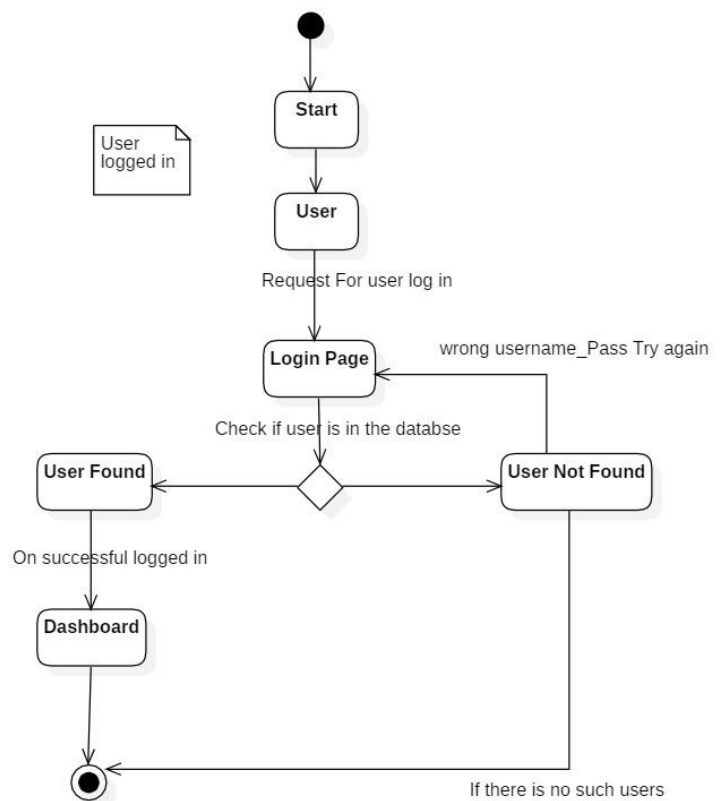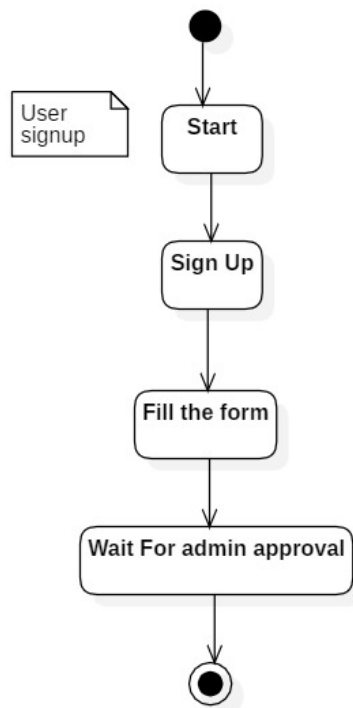

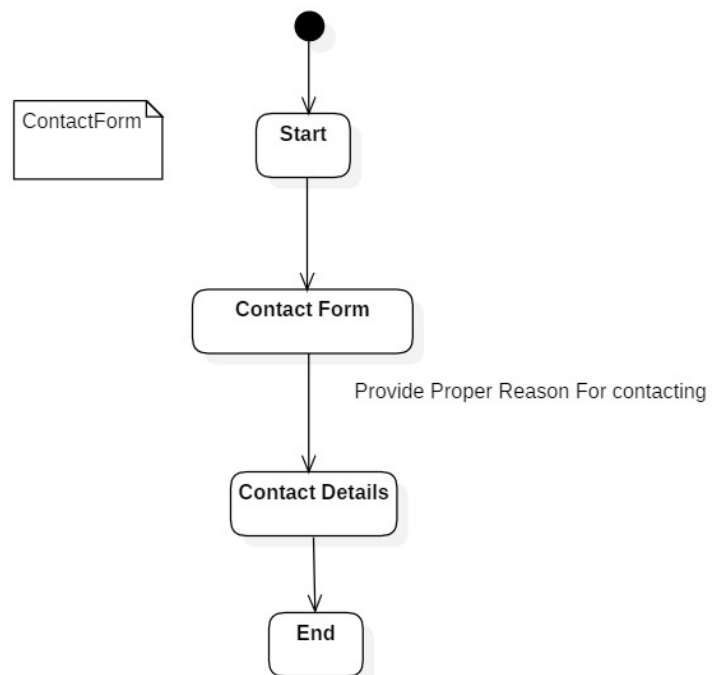*Figure 16 User Logging In*

Figure 18 User self-Sign Up

Figure 19 Form to contact for new users

Use Case Diagram
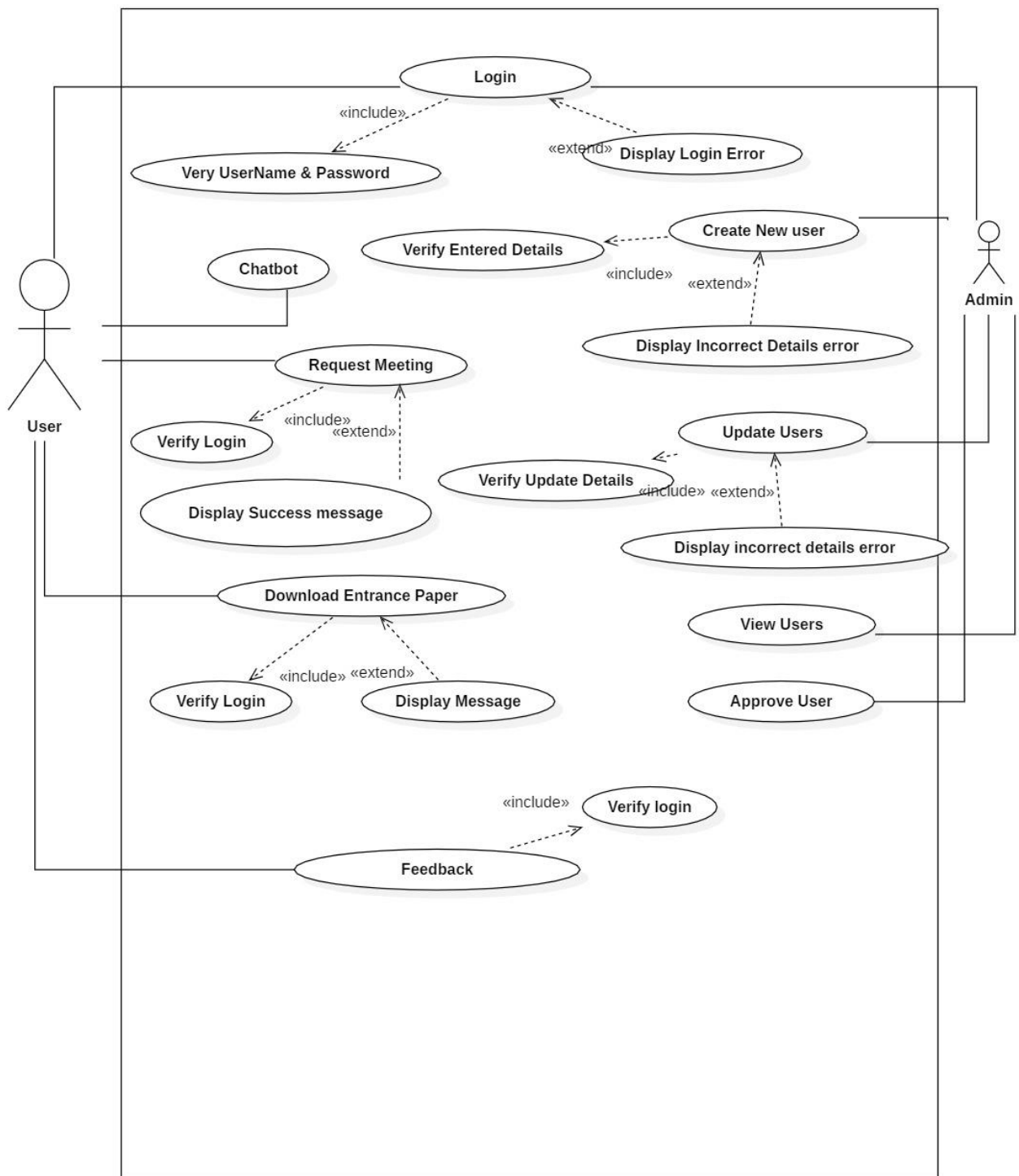
Figure 20 Use Case Diagram
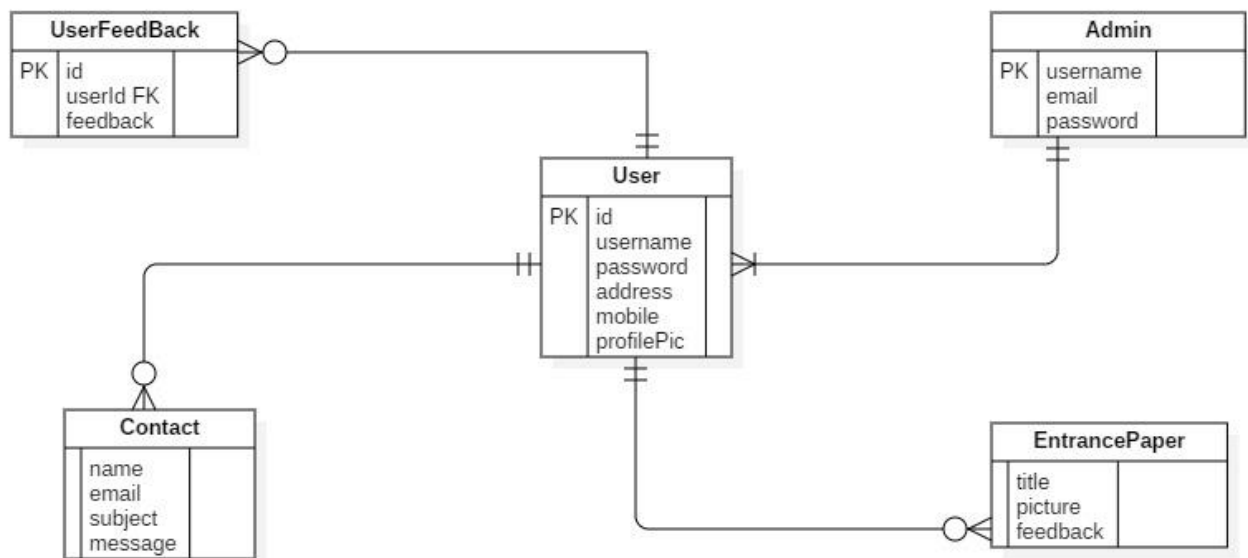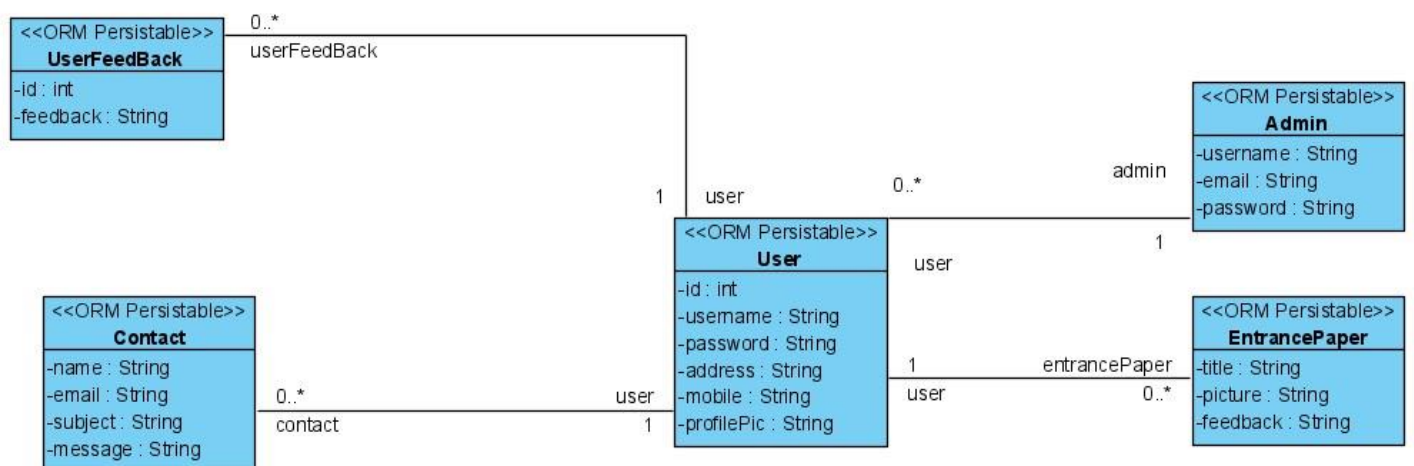
ERD Diagram

UserFeedBack

PK	id
	userId FK
	feedback

Admin

PK	username
	email
	password

User

PK	id
	username
	password
	address
	mobile
	profilePic

Contact

name
email
subject
message

EntrancePaper

title
picture
feedback

Class Diagram

<<ORM Persistable>>
UserFeedBack
- -id : int
- -feedback : String

<<ORM Persistable>>
Admin
- -username : String
- -email : String
- -password : String

<<ORM Persistable>>
User
- -id : int
- -username : String
- -password : String
- -address : String
- -mobile : String
- -profilePic : String

<<ORM Persistable>>
Contact
- -name : String
- -email : String
- -subject : String
- -message : String

<<ORM Persistable>>
EntrancePaper
- -title : String
- -picture : String
- -feedback : String

0..* userFeedBack
1 user
0..* admin
1
0..* user
contact 1
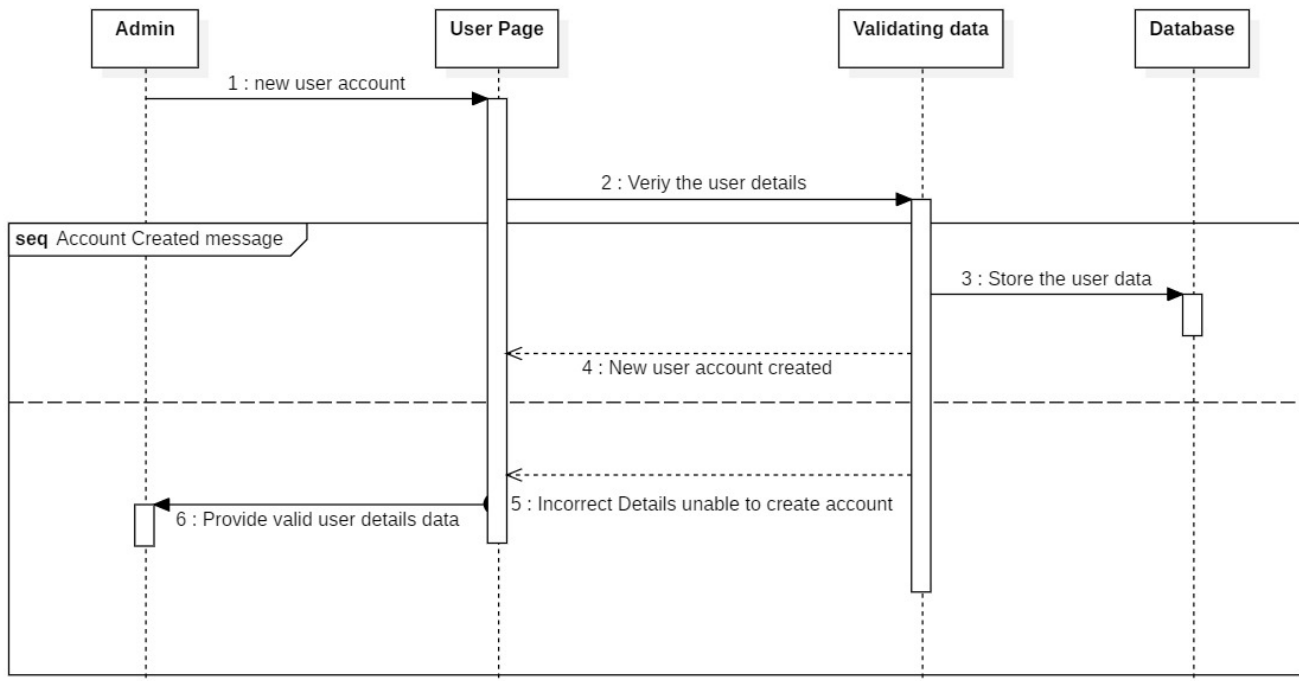1 entrancePaper
user 0..*

Sequence Diagram

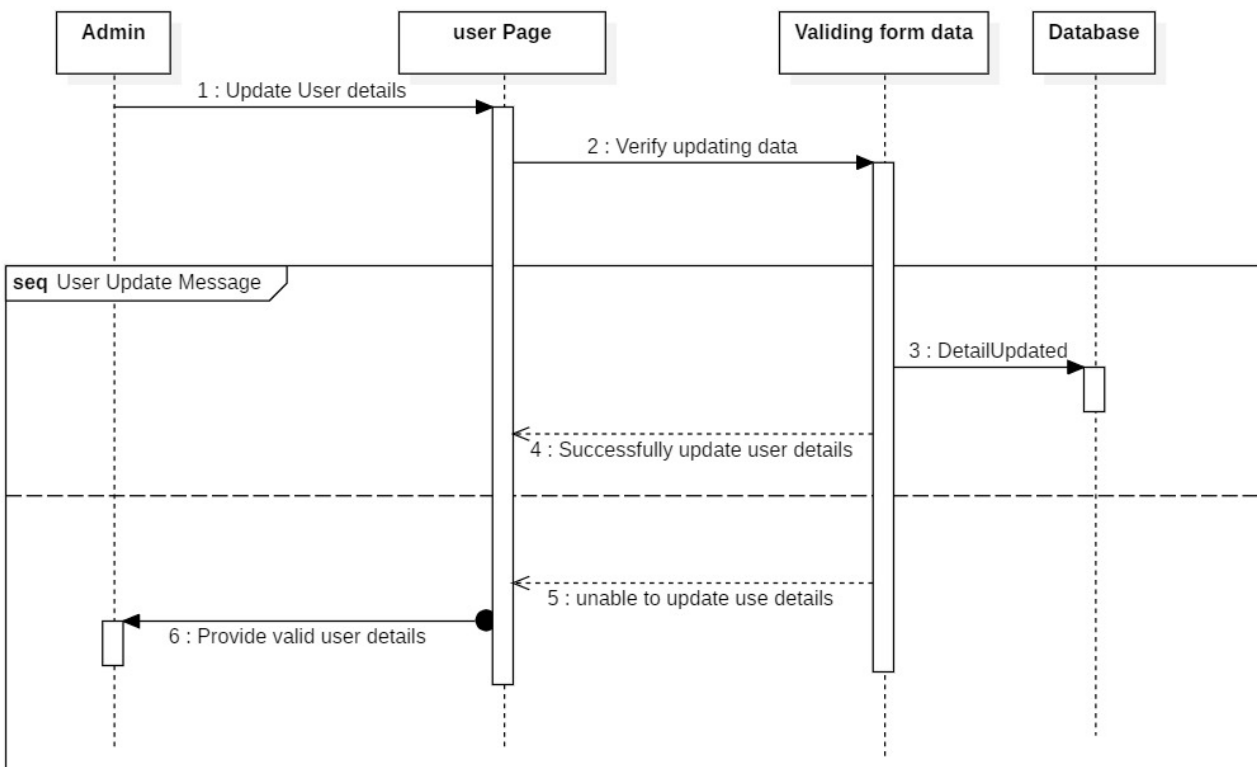*Figure 22 Creating New User Account*

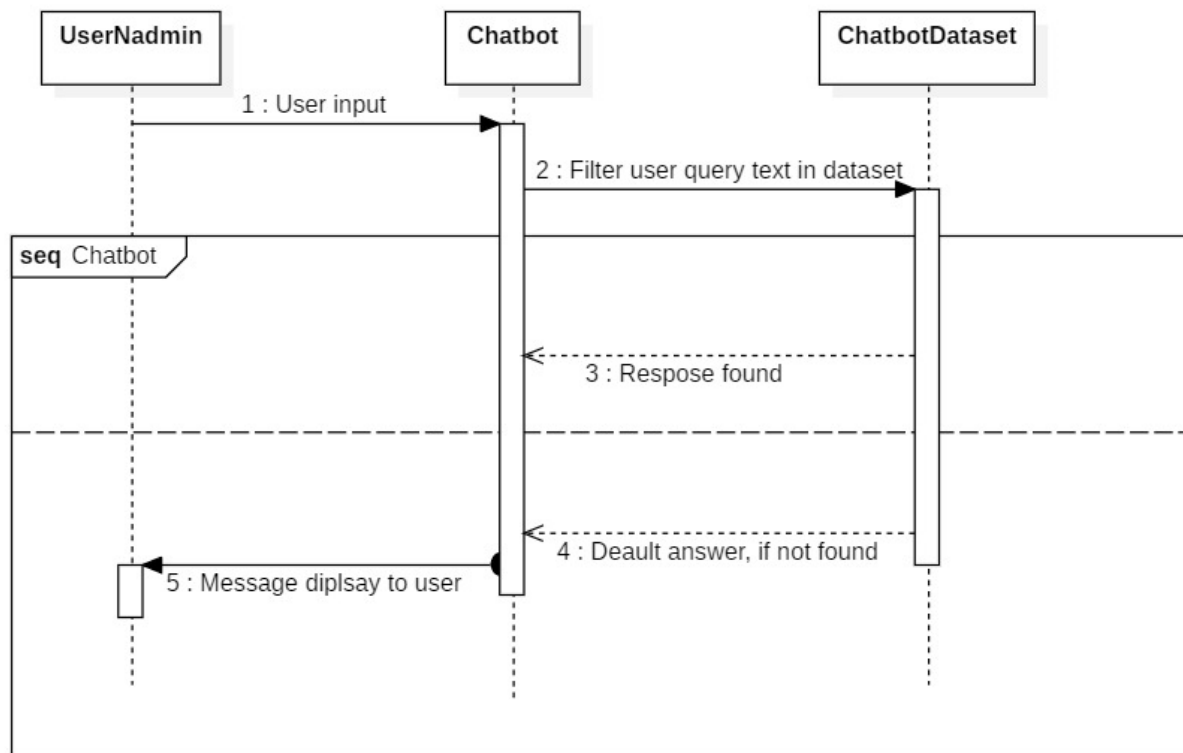

Figure 21 Updating User Details
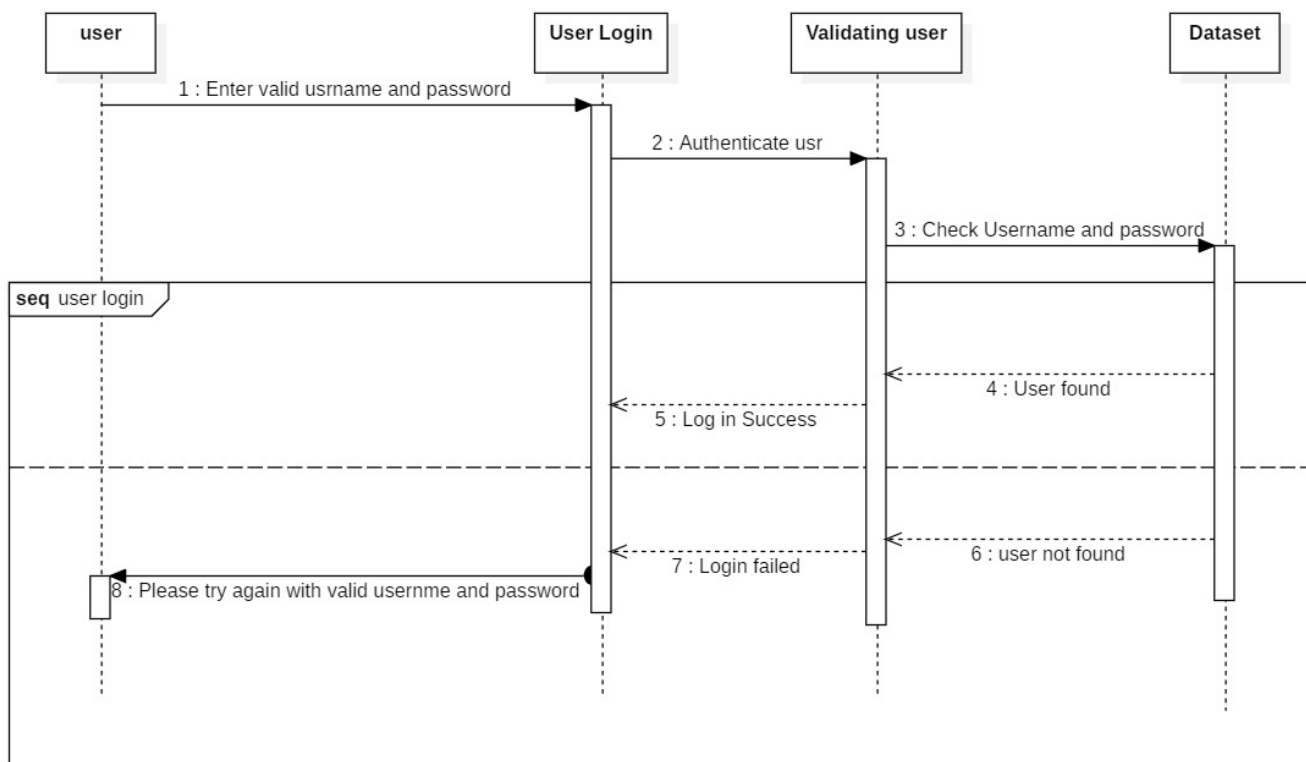
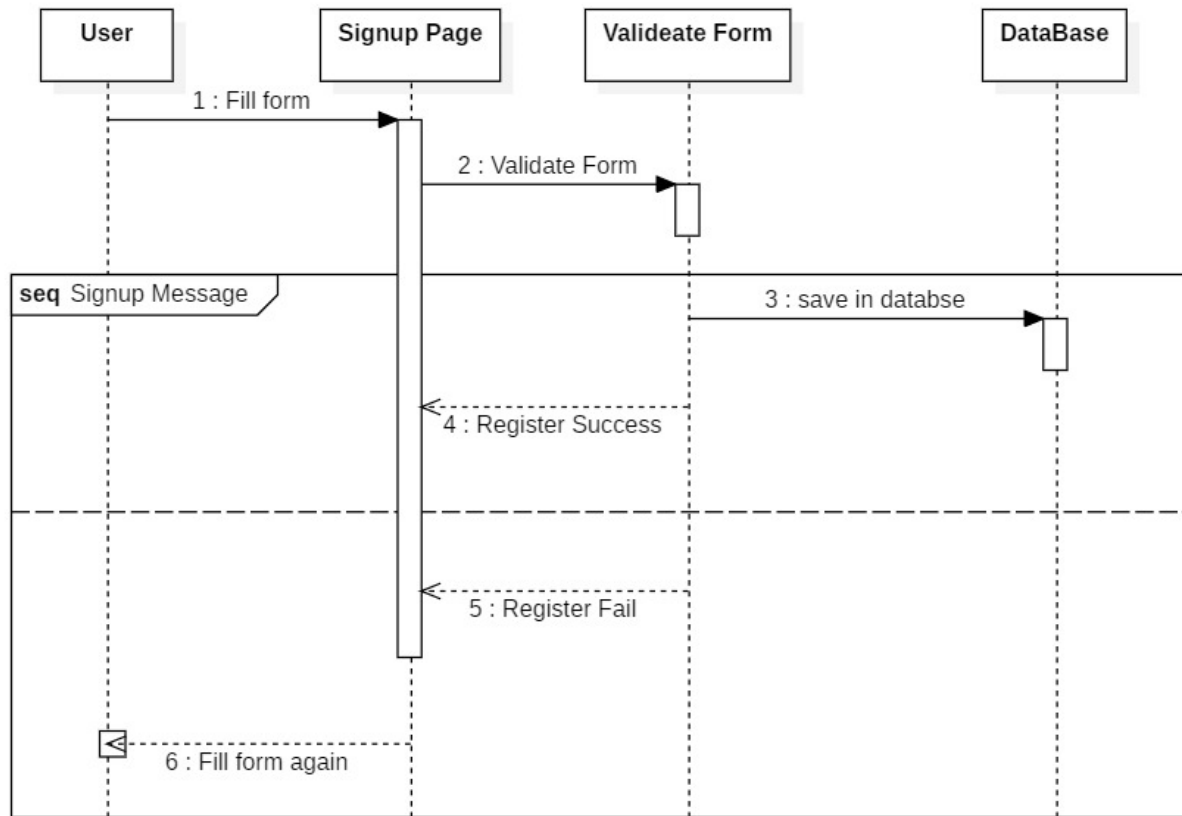*Figure 24 Chatbot Interactions*

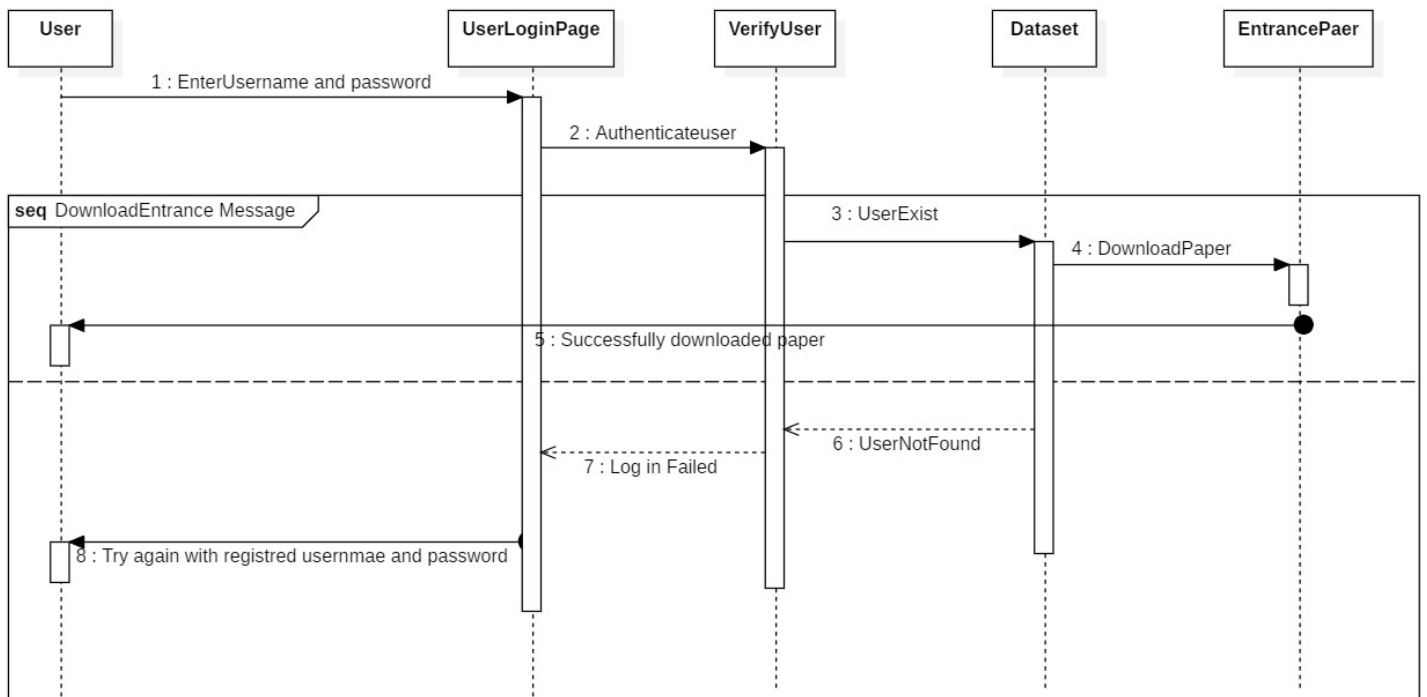

Figure 23 User logging IN

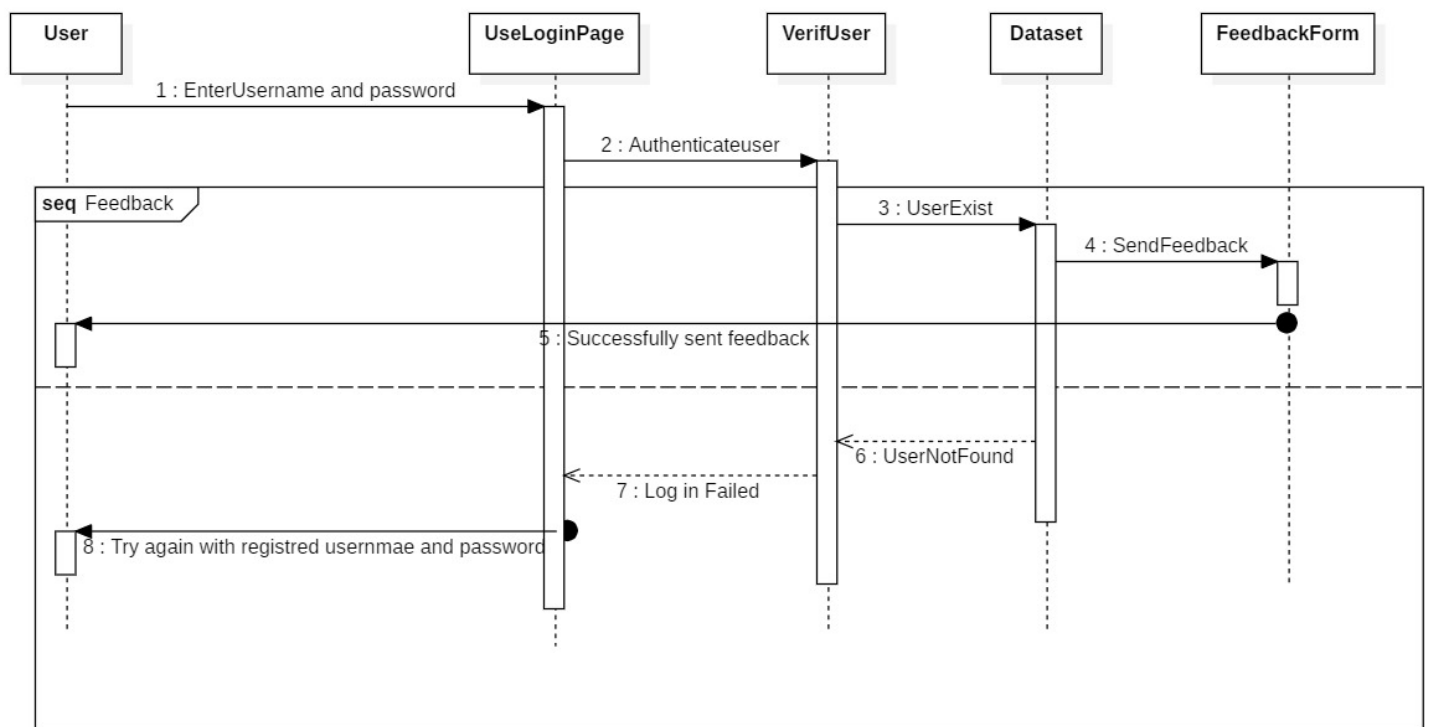*Figure 25 User Self Sign Up*

Figure 26 Entrance Paper Download

Figure 27 Feedback Form

Wireframe
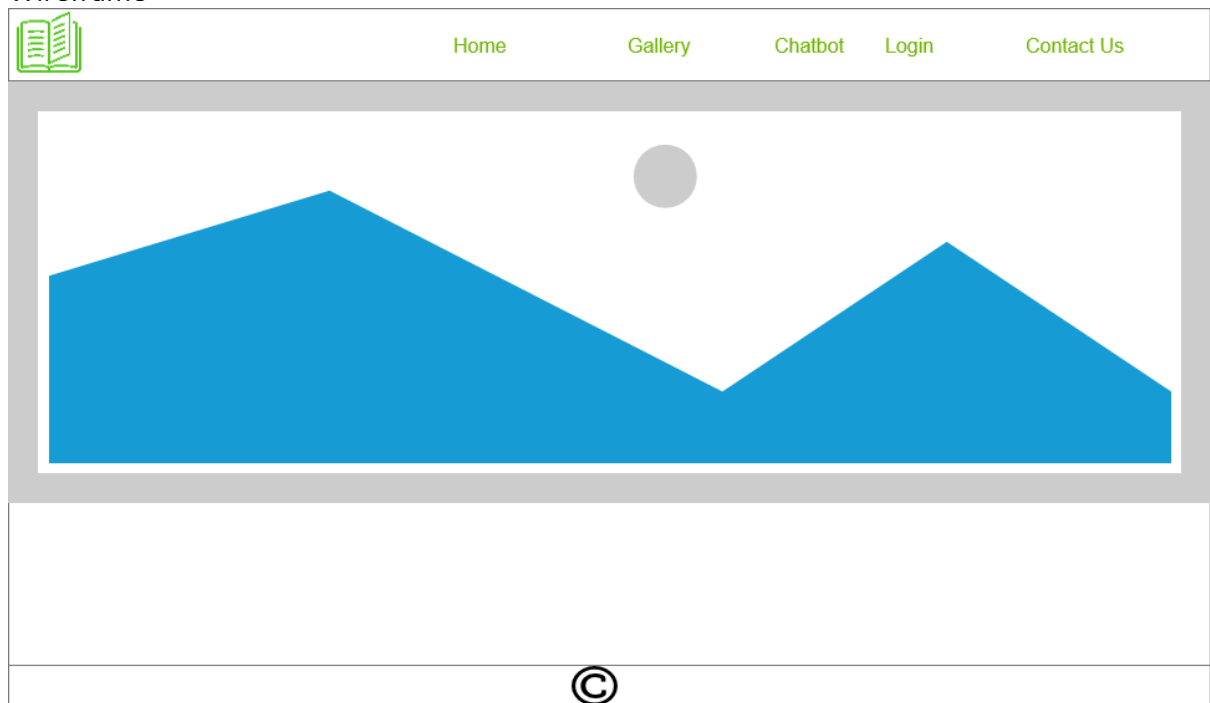
Figure 28 Landing Page

	Home	Gallery	Chatbot	Login	Contact Us

Login Form

Username/Email

Password

login

Figure 29 Login Page

	Home	Gallery	Chatbot	Login	Contact Us

SignUp Form

First Name

last name

Address

Mobile

username

Password

Choose File

Create

Figure 30 Signup Page

Figure 31 user Dash board
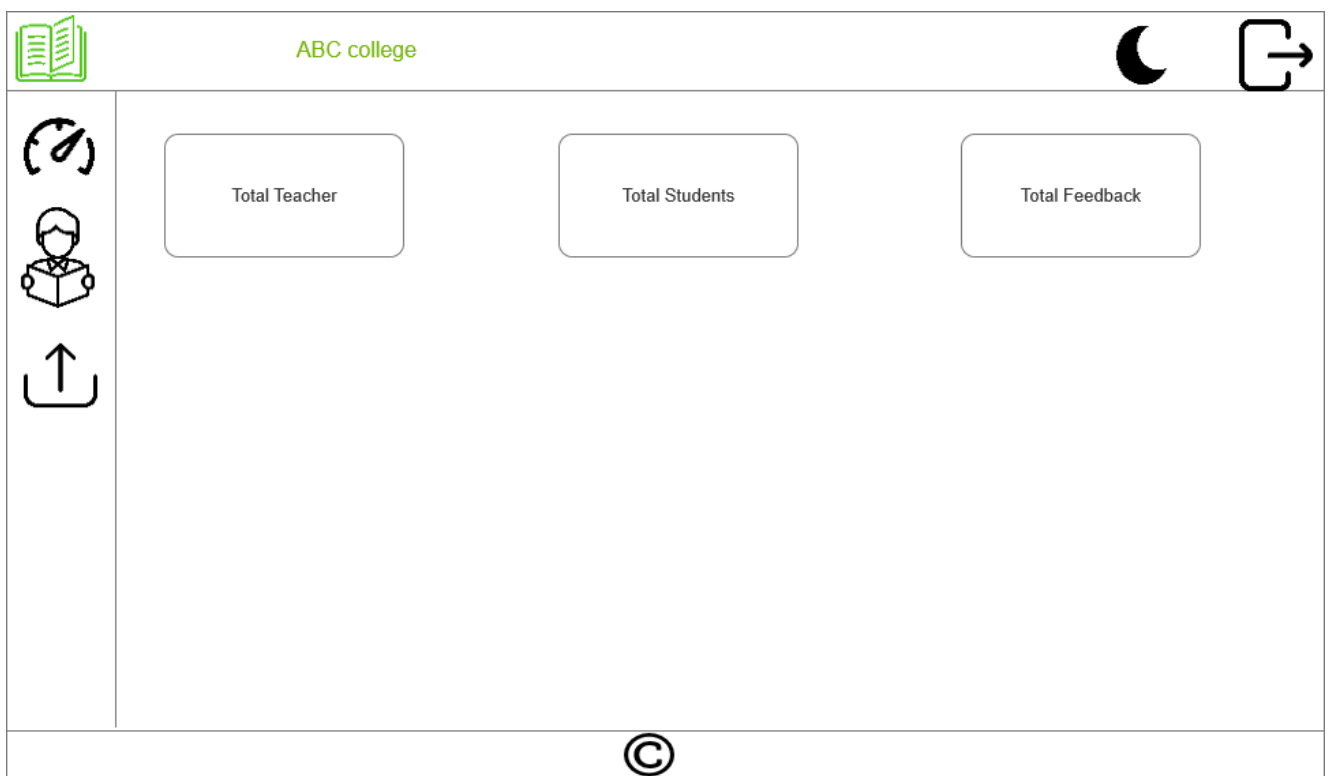
Figure 32 Admin Dashboard
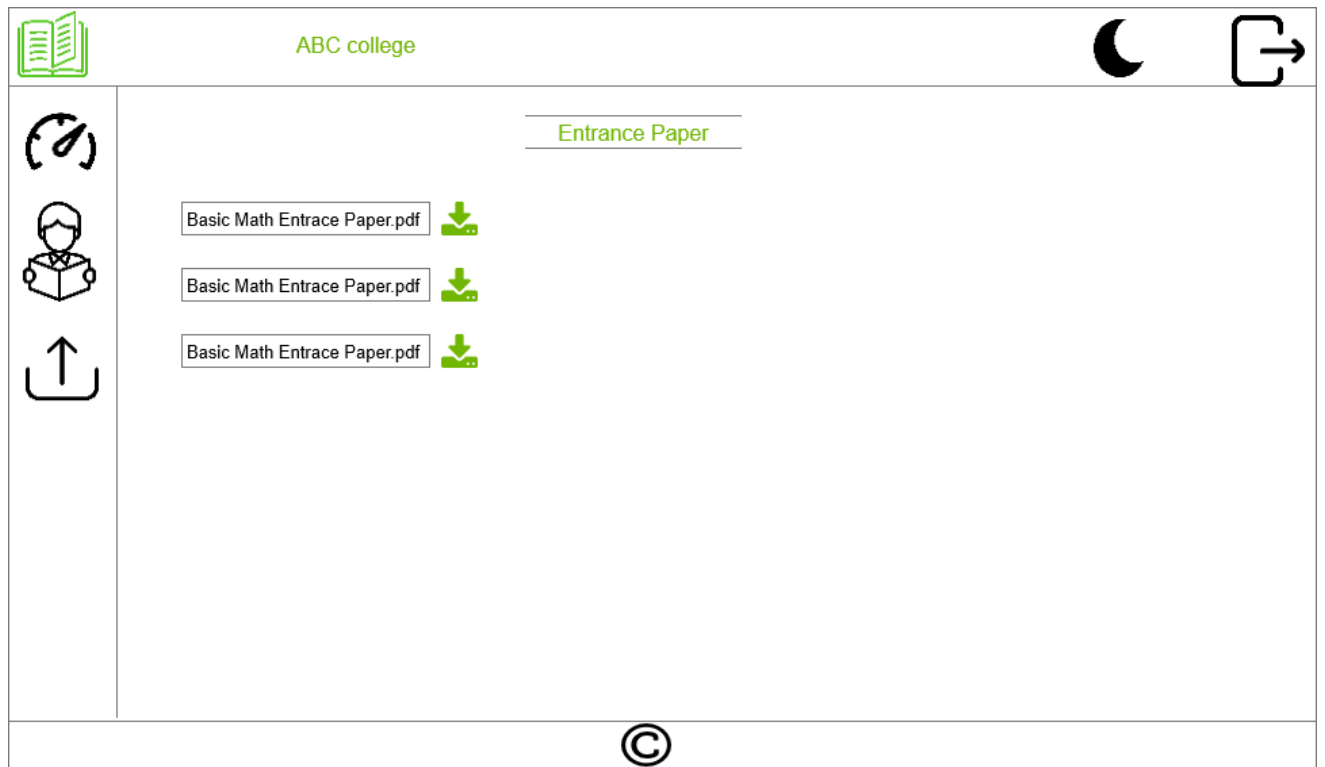
Figure 33 Entrance Paper
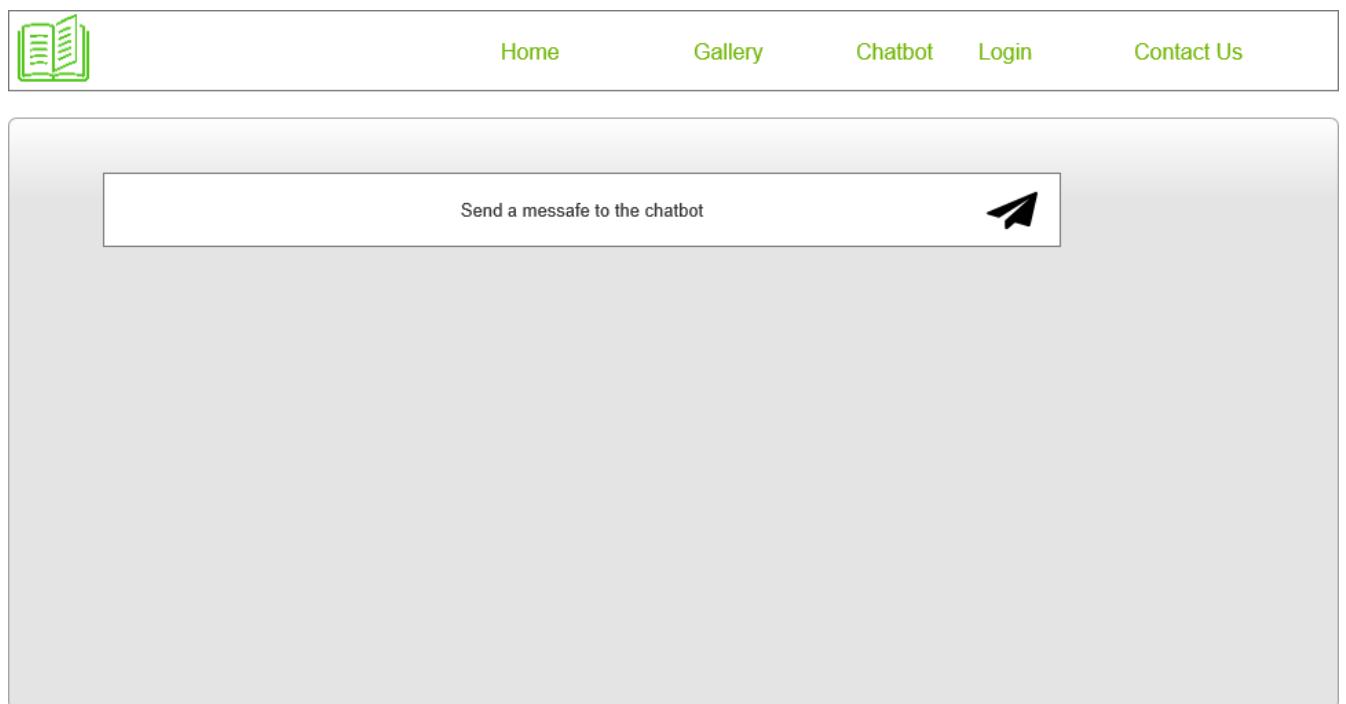
Figure 34 Chatbot GUI
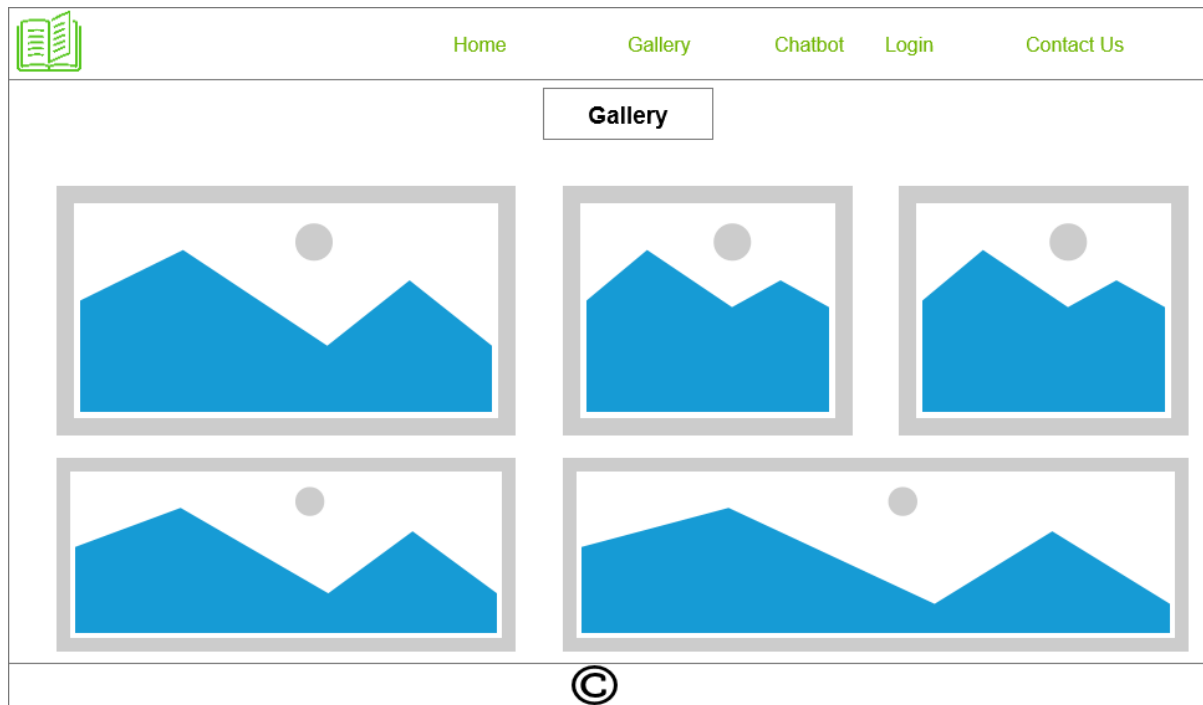
Figure 35 Gallery

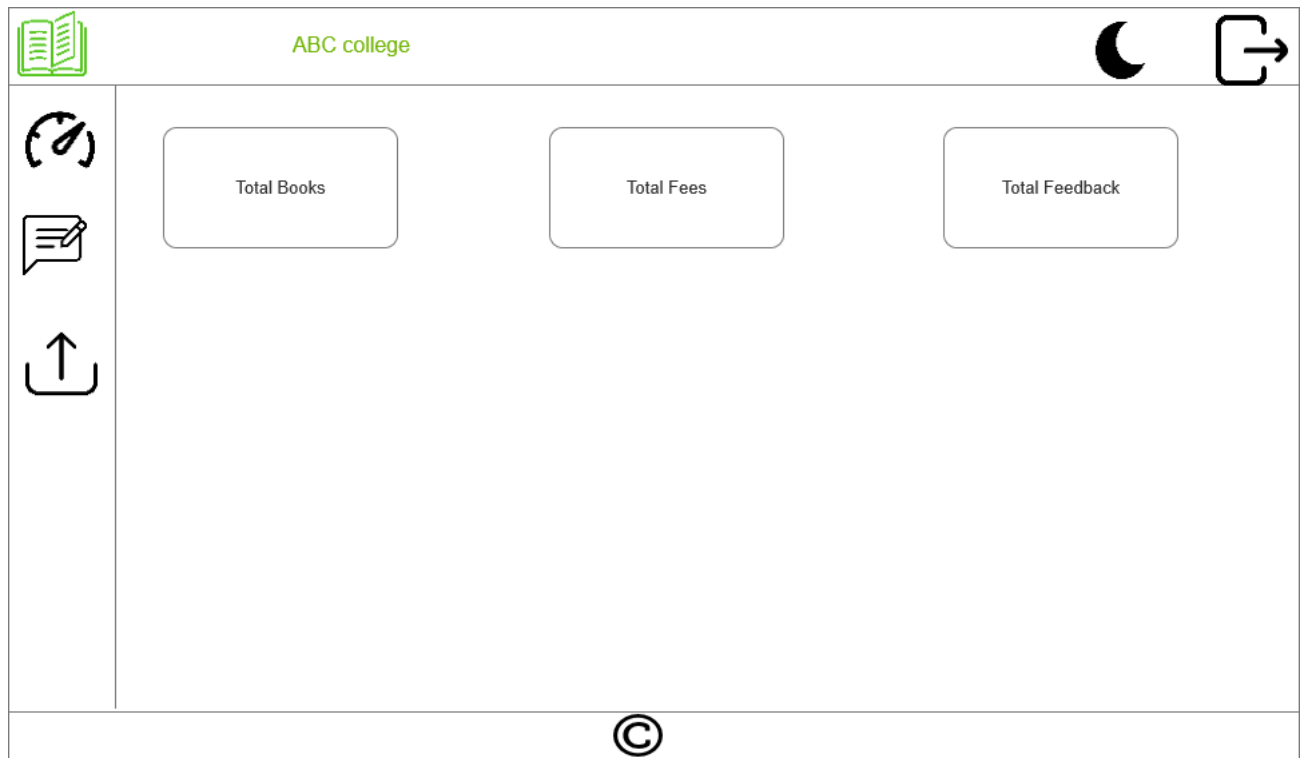## 5.3 Testing

5.3.1 Black Box Testing

It is a technique where it has been checked without coding as Junit by the features of a program. Based on functional criteria and speciation, Black Box focuses primarily on the input and performance of the software program. We can even call it as Behavioral Testing (Guru99, 2020). Following Steps are followed at testing:

a) The requirements and Specification of the systema are examined first.

b) Valid and invalid inputs are given at different scenario to detect the negative and positive scenario.

c) Tester will determine the expected of all the inputs

d) Tester will be constructing the selected inputs.

e) Test cases are executed

f) Comparison will be done between the actual output and expected output

g) Any defects will be fixed and re-test will be done [**Error! Bookmark not defined.**].

*Figure 36 Black box testing workflow [**Error! Bookmark not defined.**]*

5.3.1.1 Types of Black Box Testing
- Functional Testing
- Non-Functional Testing
- Regression Testing

5.3.1.2 Tools for Black Box Testing
- For Functional/Regression: QTP/Selenium
- For Non-Functional: LoadRunner, Jmeter

5.3.1.3 Black Box Testing Technique
- Equivalence class testing
- Boundary value testing
- Decision table testing

5.4 Plan/Schedule
5.4.1 WBS

Figure 37 Work break Down Structure

5.4.2 Gantt Chart

Figure 33 Gantt chart with Milestone

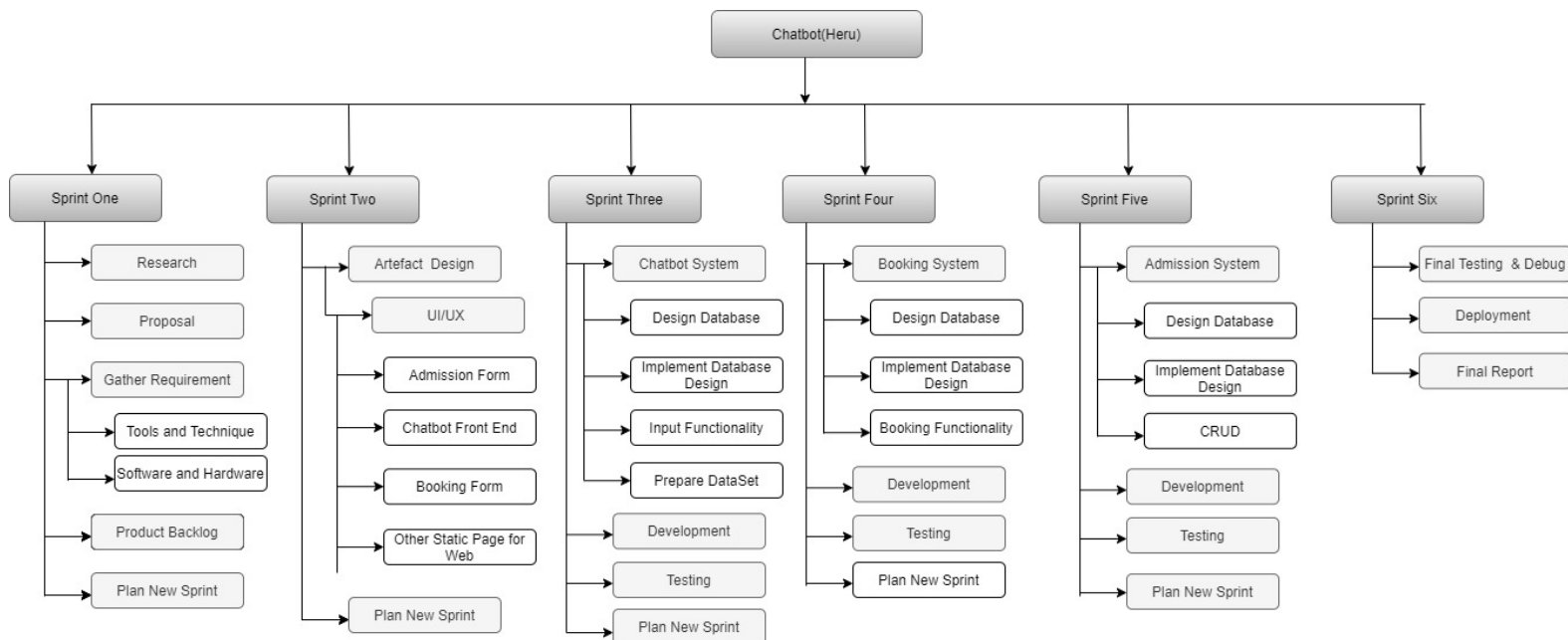| Chatbot(Heru)	0h	7%
Sprint One	**0h**	**63%**
Research	0	100%
Proposal Submission	0	100%
Gather Requirement	**0h**	**55%**
Tools and Technique	0	50%
Software & Hardware	0	60%
Product Backlog	0	60%
Literature Review Submission	0	0%
Sprint Two	**0h**	**0%**
Artefcat Design	**0h**	**0%**
UI/UX	**0h**	**0%**
Admission Form	0	0%
Chatbot Front End	0	0%
Meeting Booking	0	0%
Other Static web page	0	0%
Sprint Three	**0h**	**0%**
Chatbot system	**0h**	**0%**
Design Database	0	0%
Implement Database Design	0	0%
Input Functionality	0	0%
Prepare Dataset	0	0%
Development	0	0%
Testing	0	0%
Test Plan Submission	0	0%
Sprint Four	**0h**	**0%**
Booking System	**0h**	**0%**
Design Database	0	0%
Implement Database Design	0	0%
Booking Functionality	0	0%
Development	0	0%
Testing	0	0%
Sprint Five	**0h**	**0%**
Admission System	**0h**	**0%**
Design Database	0	0%
Implement Database Design	0	0%
CRUD functionality	0	0%
Development	0	0%
Testing	0	0%
Professionalism Submission	0	0%
Sprint Six	**0h**	**0%**
Final Testing & Debug	0	0%
Risk Analysis	0	0%
System Deployment	0	0%
Final Report Submissions (Draft)	0	0%

5.References and Bibliography

Bani, B. S. & Singh, A. P., 2017. College Enquiry Chatbot Using A.L.C.E. *International Journal of New Technology and Research,* Issue 1, pp. 64-65.

Dole, M. A., Sansare, M., Harekar, M. & Mane, P., 2016. Intelligent Chat Bot for Banking System. *International Journal for Research in Applied Science & Engineering,* 4(IV).

E., H., M.A. , B. & P., H.-y., 2009. Using scrum in global software development: a systematic literature review. *In 2009 Fourth IEEE International Conference on Global Software Engineering ,* Issue IEEE, pp. 175-184.

Gunther Cox Revision, 2019. *Chatterbot.readthedocs.io.* [Online] Available at: https://chatterbot.readthedocs.io/en/stable/ [Accessed 15 09 2020].

Guru99, 2020. *Guru99.com.* [Online] Available at: https://www.guru99.com/black-box-testing.html [Accessed 22 09 2020].

Jeremie, V. S., Jeremiah, Z. P., Leain, J. P. & Holden, J. T., 2018. A Compartive Analysis of Algorith for a School Based Chabot Using Natural Langauge processing in Text Anlysis. *CoureHero.*

J, S., A, V., J, B. & N, P., 2007. Distributed scrum: Agile project management with outsourced development teams. *40th Annual Hawaii International Conference on,* pp. 274a-2271.

P.Patel, N., R.Parikh, D., A.Patel, D. & R.Patel, R., 2019. AI and Web-Based Human-Like Interactive University Chatbot (UNIBOT). *2019 3rd International conference on Electronics, Communication and Aerospace Technology (ICECA),* pp. 148-150.

Scrum, 2020. *Scrum.Org.* [Online] Available at: https://www.scrum.org/resources/what-is-scrum [Accessed 22 09 2020].

Singh, R. et al., 2018. Chatbot using TensorFlow for small Businesses. *2018 Second International Conference on Inventive Communication and Computational Technologies (ICICCT),* pp. 1614-1619.

Deng, H., Xie, X. & Zhang, X., 2019. Enhanced Matching Network for Multi-turn Response Selection in Retrieval-Based Chatbots. 2019 International Joint Conference on Neural Networks (IJCNN), pp. 1-7.

Helmi Setyawan, M. Y., Awangga, R. M. & Efendi, S. R., 2018. Comparison Of Multinomial Naive Bayes Algorithm And Logistic Regression For Intent Classification In Chatbot. 2018 International Conference on Applied Engineering (ICAE), pp. 1-5.

Wijaya, R. A., Kusumaningtyas, E. M. & Barakbah, A., 2019. Knowledge Based Chatbot with Context Recognition. 2019 International Electronics Symposium (IES), pp. 432-438.

Yu, C., Jiang, W., Zhu, D. & Li, R., 2019. Stacked Multi-head Attention for Multi-turn Response Selection in Retrieval-based Chatbots. 2019 Chinese Automation Congress (CAC), pp. 3918-3921