

## **Abstract**

A chatbot also known as a talkbot, chatterbot, Bot, chatterbox, Artificial Conversational Entity is a computer program which conducts a conversation via auditory or textual methods. Such programs are often designed to convincingly simulate how a human would behave as a conversational partner, thereby passing the Turing Test. Chatbots are typically used in dialog systems for various practical purposes including customer service or information acquisition. Some chatterbots use sophisticated natural language processing systems, but many simpler systems scan for keywords within the input, then pull a reply with the most matching keywords, or the most similar wording pattern, from a database. Currently chatbots are widely used as part of instant messaging platforms like Facebook WeChat, and Kik for entertaining purposes as well as B2C marketing and customer service. Companies like Domino's, Pizza Hut, Disney, Nerdify, Yamatos Line and whole foods have launched their own chatbots to increase end customer engagement, promote their products and services, and give their customers a more convenient and easier way to order from them. In 2016 in the travel industry, several agencies and airlines launched chatbot services via Messenger Aeromexico's sells tickets and answers questions using artificial intelligence, and both Aeromexico's and KLM's provide status updates, allow users to check in for deliver mobile boarding passes and recommend hotels, restaurants and things to do in the destination. Chinese travel companies had already been providing these services for some time via WeChat.

# INDEX

<b>1 Synopsis</b>	<b>1</b>
1.1 Project Title . . . . .	2
1.2 Project Option . . . . .	2
1.3 Internal Guide . . . . .	2
1.4 Sponsorship and External Guide . . . . .	2
1.5 Technical Keywords . . . . .	2
1.6 Problem Statement . . . . .	2
1.7 Abstract . . . . .	3
1.8 Goals and Objectives . . . . .	3
1.9 Relevant mathematics associated with the Project . . . . .	3
1.10 Names of Conferences / Journals where papers can be published . .	4
1.11 Review of Conference/Journal Papers supporting Project idea . . .	4
1.12 Plan of Project Execution . . . . .	6
<b>2 Technical Keywords</b>	<b>7</b>
2.1 Area of Project . . . . .	8
2.2 Technical Keywords . . . . .	8
<b>3 Introduction</b>	<b>9</b>
3.1 Project Idea . . . . .	10
3.2 Motivation of the Project . . . . .	10
3.3 Literature Survey . . . . .	11
<b>4 Problem Definition and Scope</b>	<b>12</b>
4.1 Problem Statement . . . . .	13

4.1.1	Goals and objectives . . . . .	13
4.2	Software context . . . . .	13
4.3	Major Constraints . . . . .	13
4.4	Methodologies of Problem solving and efficiency issues . . . . .	14
4.5	Outcome . . . . .	14
4.6	Applications . . . . .	14
4.7	Hardware Resources Required . . . . .	14
4.8	Software Resources Required . . . . .	14
<b>5</b>	<b>Project Plan</b>	<b>16</b>
5.1	Project Estimates . . . . .	17
5.1.1	Reconciled Estimates . . . . .	18
5.1.2	Project Resources . . . . .	19
5.2	Risk Management w.r.t. NP Hard analysis . . . . .	19
5.2.1	Risk Identification . . . . .	19
5.2.2	Risk Analysis . . . . .	19
5.3	Project Schedule . . . . .	19
5.3.1	Project task set . . . . .	19
5.3.2	Task Network . . . . .	22
5.3.3	Timeline Chart . . . . .	22
5.4	Team Organization . . . . .	22
5.4.1	Team structure . . . . .	22
5.4.2	Management reporting and communication . . . . .	23
<b>6</b>	<b>Software Requirement Specification</b>	<b>24</b>
6.1	Introduction . . . . .	25
6.1.1	Purpose and Scope of Document . . . . .	25
6.1.2	Overview of responsibilities of Developer . . . . .	25
6.2	Usage Scenario . . . . .	26
6.2.1	Use-cases . . . . .	26
6.2.2	Use Case View . . . . .	27
6.3	Data Model and Description . . . . .	27

6.3.1	Data Description . . . . .	27
6.3.2	Data objects and Relationships . . . . .	28
6.4	Functional Model and Description . . . . .	29
6.4.1	Data Flow Diagram . . . . .	30
6.4.2	Activity Diagram: . . . . .	31
6.4.3	Non Functional Requirements: . . . . .	31
6.4.4	State Diagram: . . . . .	32
6.4.5	Design Constraints . . . . .	32
6.4.6	Software Interface Description . . . . .	33
<b>7</b>	<b>Detailed Design Document</b>	<b>34</b>
7.1	Introduction . . . . .	35
7.2	Architectural Design . . . . .	35
7.3	Compoent Design . . . . .	36
7.3.1	Class Diagram . . . . .	36
<b>8</b>	<b>Project Implementation</b>	<b>37</b>
8.1	Introduction . . . . .	38
8.2	Tools and Technologies Used . . . . .	38
8.2.1	Heroku : . . . . .	39
8.3	Methodologies/Algorithm Details . . . . .	39
8.3.1	Algorithm . . . . .	39
<b>9</b>	<b>Software Testing</b>	<b>41</b>
9.1	Type of Testing Used . . . . .	42
9.2	Testing . . . . .	42
9.3	Test Cases and Test Results . . . . .	46
<b>10</b>	<b>Results</b>	<b>48</b>
10.1	Outputs . . . . .	49
<b>11</b>	<b>Deployment and Maintenance</b>	<b>52</b>
11.1	Installations: . . . . .	53

<b>12 Conclusion and Future Scope</b>	<b>54</b>
12.1 Conclusion: . . . . .	55
12.2 Future Scope: . . . . .	55
<b>13 References</b>	<b>56</b>
<b>Annexure A Laboratory assignments on Project Analysis of Algorithmic Design</b>	<b>59</b>
<b>Annexure B Project Planner</b>	<b>62</b>
<b>Annexure C Reviewers Comments of Paper Submitted</b>	<b>64</b>
C.1 Review of Conference/Journal Papers supporting Project idea . . . . .	65
<b>Annexure D Plagiarism Report</b>	<b>66</b>
<b>Annexure E Term-II Project Laboratory Assignments</b>	<b>68</b>
<b>Annexure F Information of Project Group Members</b>	<b>70</b>

# List of Figures

1.1	Project Plan . . . . .	6
5.1	Task Network . . . . .	22
6.1	Use case diagram . . . . .	28
6.2	Entity Relationship Diagram . . . . .	28
6.3	Class Diagram . . . . .	29
6.4	Data Flow Diagram 0 . . . . .	30
6.5	Data Flow Diagram 1 . . . . .	30
6.6	Activity diagram . . . . .	31
6.7	State Transition Diagram . . . . .	32
7.1	Architecture Diagram . . . . .	35
7.2	Class Diagram . . . . .	36
10.1	Communication regarding admission enquiry . . . . .	49
10.2	Communication regarding college information . . . . .	50
10.3	Communication regarding greeting . . . . .	51
A.1	Canvas Graph . . . . .	61
B.1	Project Plan . . . . .	63

# List of Tables

4.1	Hardware Requirements . . . . .	14
5.1	Time Estimate . . . . .	18
5.2	Risk Table . . . . .	19
5.3	Risk Probability definitions . . . . .	20
5.4	Risk Probability definitions . . . . .	20
5.5	Risk Impact definitions . . . . .	20
5.6	Team Responsibilities . . . . .	23
6.1	Use Cases . . . . .	27
9.1	Test Cases . . . . .	46
9.2	Test Cases . . . . .	47
A.1	IDEA Matrix . . . . .	60
A.2	Test Cases for Black Box testing . . . . .	61

## **CHAPTER 1**

### **SYNOPSIS**

## **1.1 PROJECT TITLE**

ChatBot for Education System

## **1.2 PROJECT OPTION**

N.A.

## **1.3 INTERNAL GUIDE**

Dr. K.S.WAGH

## **1.4 SPONSORSHIP AND EXTERNAL GUIDE**

N.A.

## **1.5 TECHNICAL KEYWORDS**

**Technical Key Words:**

- Artificial Intelligence
- Keyword Extraction
- Data Mining
- Pattern Matching
- Machine Learning

## **1.6 PROBLEM STATEMENT**

To develop a system which will give automatic reply on behalf of human. The user will type a query and immediately he or she will get a reply. This will reduce the overhead of user as well because one doesn't need to call the institute for information, or doesn't need to download the huge document and go through it just to search a single information.

## **1.7 ABSTRACT**

A chatbot also known as a talkbot, chatterbot, Bot, chatterbox, Artificial Conversational Entity is a computer program which conducts a conversation via auditory or textual methods. Such programs are often designed to convincingly simulate how a human would behave as a conversational partner, thereby passing the Turing Test. Chatbots are typically used in dialog systems for various practical purposes including customer service or information acquisition. Some chatterbots use sophisticated natural language processing systems, but many simpler systems scan for keywords within the input, then pull a reply with the most matching keywords, or the most similar wording pattern, from a database. Currently chatbots are widely used as part of instant messaging platforms like Facebook WeChat, and Kik for entertaining purposes as well as B2C marketing and customer service. Companies like Domino's, Pizza Hut, Disney, Nerdify, Yamatos Line and whole foods have launched their own chatbots to increase end customer engagement, promote their products and services, and give their customers a more convenient and easier way to order from them. In 2016 in the travel industry, several agencies and airlines launched chatbot services via Messenger Aeromexico's sells tickets and answers questions using artificial intelligence, and both Aeromexico's and KLM's provide status updates, allow users to check in for deliver mobile boarding passes and recommend hotels, restaurants and things to do in the destination. Chinese travel companies had already been providing these services for some time via WeChat.

## **1.8 GOALS AND OBJECTIVES**

- The goal of our project is to remove the human intervention.
- To develop reliable and highly accurate automatic response generating system.
- To make interaction very quick between user and system.

## **1.9 RELEVANT MATHEMATICS ASSOCIATED WITH THE PROJECT**

System Description: { I,O,S,F }

- Chat Bot Input:

$I = \{ A_1, A_2, A_3, \dots, A_n \}$

- Chat Bot Output:

$O = O_1, O_2, \dots, O_n$

Chat Bot will give appropriate reply

- Success Conditions:

$S = \{ \text{If chat bot gives relevant reply to user's query.} \}$

- Failure Conditions:

$F = \{ \text{If Chat Bot gives irrelevant answers to user's query.} \}$

## **1.10 NAMES OF CONFERENCES / JOURNALS WHERE PAPERS CAN BE PUBLISHED**

- IJETCS

## **1.11 REVIEW OF CONFERENCE/JOURNAL PAPERS SUPPORTING PROJECT IDEA**

- Kyo-Joong oh et, al in [1] have proposed a chat bot psychiatric counseling in mental health care service. Emotion recognition is used to implement this chat bot. Because of this , patient is not needed at all to go to hospital. Free counseling is provided at users location.
- Nikita Hatwar in [2] proposed another interesting chat bot for marketing eld named as AI based chat bot. In this paper , this chat bot is used for providing guidance to the visitor of mall, provides navigation and according to latest discount going in the shop, it was guiding recommendation system is implemented in this project according to the rating we can advice people to go to particular shop. This feature can be extended further.
- Yu Wo and Gongxiao Wang in [3] presented paper automatic chat bot knowledge acquisition from online via rough set and ensemble learning. The concept

of ensemble learning is used here. The which construct classification results of the learner to get the final result. Multiple rough set classifier are constructed and trained first , then all replies are classified with these classifiers . the final results are drawn by voting to the output of these classifiers and finally it is selected as knowledge database . the disadvantage of this system is that not all replies are related to root message different forum have different styles and formats.

- Linga chandar P.C in [4] have proposed an android based educational chat bot for visually impaired people. Voice processing ,symbolic reduction and keyword detection methodologies are used. In case of noisy environment application is not recommended. Application should be scalable , so even if thousands of queries are asked at a time. Then also system is able to answer all the queries.
- Juan carlos, et al[5] proposed system that is chatting with arduino platform through telegram bot. The proposed system is connecting telegram and arduino platform using telegram bot .this system allows person to communicate with machine user, where this machine user can operate simple and complex prototype developed by same person. To implement this system,telegram api are used .REST protocol is used for the integration of both.
- Jacques Gracia Fausto abraham in [6] This paper describes how to implement chat bot on twitter for entertainment and viral advertising using database and simple algorithms. Message reception, message passing, message processing techniques are used. With this system a person can earn more followers everyday without any human intervention.
- Wenzeng Zhu in [7] proposed a system that is internet security protection for IRC based BotNet many works contain infected data which can harm our system. So in this paper, they proposed a noble paper based on IRC analysis and detection to protect our computers. IRC servers used instead of CC server. So the infected computers link to the IRC server. This method can protect the infected computers not to return Bots controlled by the Bot Master. The only problem with this is that we need to detect all the infected computers and help

them to uninstall bot code. To return security, but problem is that some Bot code are hard to uninstall.

## 1.12 PLAN OF PROJECT EXECUTION

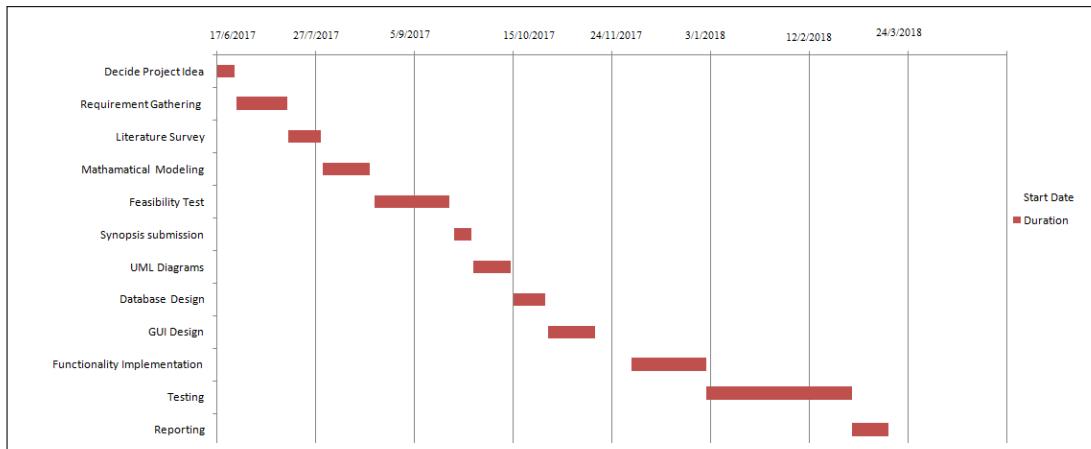


Figure 1.1: Project Plan

## **CHAPTER 2**

## **TECHNICAL KEYWORDS**

## **2.1 AREA OF PROJECT**

- Artificial Intelligence

## **2.2 TECHNICAL KEYWORDS**

### **Technical Key Words:**

- Artificial Intelligence
- Machine Learning
- Keyword Extraction
- Pattern Matching
- Data Mining

## **CHAPTER 3**

### **INTRODUCTION**

### **3.1 PROJECT IDEA**

To develop a system which will give automatic reply on behalf of human. The user will type a query and immediately he or she will get a reply. This will reduce the overhead of user as well because one doesn't need to call the institute for information, or doesn't need to download the huge document and go through it just to search a single information.

### **3.2 MOTIVATION OF THE PROJECT**

- Need a lot of searching while taking admission to the college. Student come to institute and or call up the institute for enquiry. Because of this there is a need of telephone operator or a person who will satisfy students questions.
- Huge amount of data is available. Going through that data is very tedious task for normal user. So this system will solve that problem.

### 3.3 LITERATURE SURVEY

Sr. No	Paper Name	Author	Published Year	Description
1	psychiatric counseling in mental health care service.	Kyo-Joong	2016	Emotion recognition is used to implement this chat bot. Because of this, the patient is not needed at all to go to hospital. Free counseling is provided at the user's location.
2	AI based chat bot	Nikita Hatwar	2016	Providing guidance to the visitor of the mall, provides navigation and according to latest discount going in the shop, it was guiding.
3	Android based educational chat bot	Nitin Kumar	2017	Voice processing ,symbolic reduction and keyword detection methodologies are used.
4	Internet security protection for IRC based BotNet	Wenzeng Zhu	2015	IRC analysis and detection to protect our computers. IRC servers, used instead of C&C server. So the infected computers link to the IRC server.

## **CHAPTER 4**

### **PROBLEM DEFINITION AND SCOPE**

## **4.1 PROBLEM STATEMENT**

To develop a system which will give automatic reply on behalf of human. The user will type a query and immediately he or she will get a reply. This will reduce the overhead of user as well because one doesn't need to call the institute for information, or doesn't need to download the huge document and go through it just to search a single information.

### **4.1.1 Goals and objectives**

- The goal of our project is to make use of an artificial intelligence technology to give automatic response to the user's query.
- The objective of the proposed system is that user doesn not need to call to the institute for enquire or does not need to download the document for searching information. This will save the time of user.

## **4.2 SOFTWARE CONTEXT**

- The application of the software can be used in a smart institute.
- Along with this there will be a registration system, which will collect the student information from them and that information can be used further for marketing purpose of college.

## **4.3 MAJOR CONSTRAINTS**

- The major constraint of this system is that it can't give all questions answer. There should be an admin system which will give reply for that question and this will be again a human intervention
- The system should be in a continuous internet connected environment.

#### **4.4 METHODOLOGIES OF PROBLEM SOLVING AND EFFICIENCY IS-SUES**

- More and more training is required:

If system is not trained properly then it may give irrelevant answers. Either that question should forward to admin system or user has to get satisfied by that irrelevant answer.

#### **4.5 OUTCOME**

- User can ask any question regarding institute and he or she will get proper answer.

#### **4.6 APPLICATIONS**

- The project can be implemented for Smart Institute.
- The data collected will be sent to the office body for sending some important notifications related to the institute.
- Searching problem can be reduced drastically.

#### **4.7 HARDWARE RESOURCES REQUIRED**

Sr. No.	Parameter	Minimum Requirement	Justification
1	CPU Speed	2 GHz	Heavy processes
2	RAM	3 GB	For Android platform

Table 4.1: Hardware Requirements

#### **4.8 SOFTWARE RESOURCES REQUIRED**

Platform :

1. Operating System: Windows,Linux

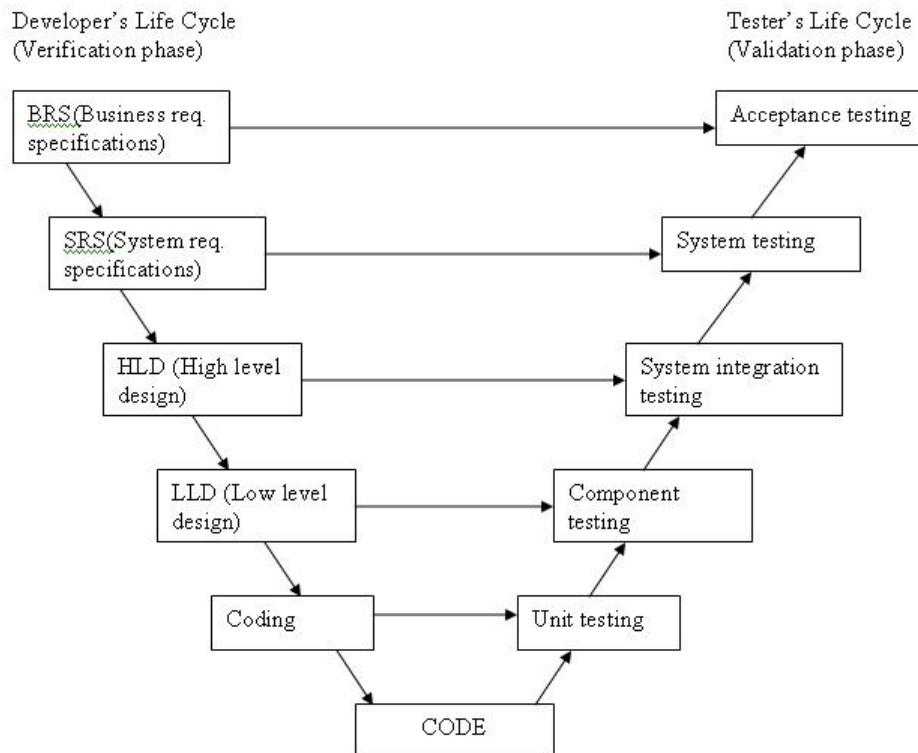
2. IDE: Eclipse IDE
3. Programming Language: Python

# **CHAPTER 5**

## **PROJECT PLAN**

## 5.1 PROJECT ESTIMATES

V- model means Verification and Validation model. Just like the waterfall model, the V-Shaped life cycle is a sequential path of execution of processes. Each phase must be completed before the next phase begins. Testing of the product is planned in parallel with a corresponding phase of development in V-model.



The various phases of the V-model are as follows:

**Requirements** like BRS and SRS begin the life cycle model just like the waterfall model. But, in this model before development is started, a system test plan is created. The test plan focuses on meeting the functionality specified in the requirements gathering.

**The high-level design (HLD)** phase focuses on system architecture and design. It provides an overview of the solution, platform, system, product and service/process. An integration test plan is created in this phase as well in order to test the pieces of the software system's ability to work together.

**The low-level design (LLD)** phase is where the actual software components are designed. It defines the actual logic for each and every component of the system. Class diagram with all the methods and relation between classes comes under LLD. Component tests are created in this phase as well.

**The implementation** phase is, again, where all coding takes place. Once coding is complete, the path of execution continues up the right side of the V where the test plans developed earlier are now put to use.

**Coding:** This is at the bottom of the V-Shape model. Module design is converted into code by developers.

### 5.1.1 Reconciled Estimates

#### 5.1.1.1 Time Estimates

Sr.No.	Task Name	Start Date	Duration	End Date
1	Decide Project Idea	17/06/2016	7	24/06/2016
2	Requirement Gathering and Testing	25/06/2016	10	15/07/2016
3	Literature Survey	16/07/2016	13	29/07/2016
4	Mathamatical Modeling	30/07/2016	19	19/08/2016
5	Planning and Feasibility Test	20/08/2016	30	20/09/2016
6	Synopsis submission	21/09/2016	7	28/09/2016
7	UML Diagrams	29/09/2016	15	14/10/2016
8	Database Design	15/10/2016	13	28/10/2016
9	GUI Design	29/10/2016	19	18/11/2016
10	Coding	02/12/2016	30	31/01/2017
11	Testing	01/02/2016	28	28/02/2017
12	Documentation	01/03/2017	15	15/03/2017

Table 5.1: Time Estimate

#### 5.1.1.2 Cost Estimate

COST:16000

### **5.1.2 Project Resources**

- **People:**

1)Internal Project Guide

- **Software:**

1)Python with Flask package.

2)ChatterBot package

## **5.2 RISK MANAGEMENT W.R.T. NP HARD ANALYSIS**

### **5.2.1 Risk Identification**

At the time of project execution proper searching of data is required. For this purpose, we need to train our system as much possible as.

### **5.2.2 Risk Analysis**

The risks for the Project can be analyzed with response giving functionality.

ID	Risk Description	Probability	Impact		
			Schedule	Quality	Overall
1	Inappropriate threshold value	Medium	Low	High	Medium
2	Inappropriate Data Base	High	High	High	High
3	Inappropriate queries	High	Low	High	High
4	Chat Bot system crash	Medium	High	High	High

Table 5.2: Risk Table

## **5.3 PROJECT SCHEDULE**

### **5.3.1 Project task set**

Major Tasks in the Project stages are:

Probability	Value	Description
High	Probability of occurrence is	> 75%
Medium	Probability of occurrence is	26 – 75%
Low	Probability of occurrence is	< 25%

Table 5.3: Risk Probability definitions

Probability	Value	Description
High	Probability of occurrence is	> 75%
Medium	Probability of occurrence is	26 – 75%
Low	Probability of occurrence is	< 25%

Table 5.4: Risk Probability definitions

Impact	Value	Description
Very high	> 10%	Schedule impact or Unacceptable quality
High	5 – 10%	Schedule impact or Some parts of the project have low quality
Medium	< 5%	Schedule impact or Barely noticeable degradation in quality Low Impact on schedule or Quality can be incorporated

Table 5.5: Risk Impact definitions

Risk ID	1
Risk Description	Inappropriate queries.
Category	Software.
Source	This was identified during early development and testing.
Probability	High
Impact	Very High
Strategy	Setting default values for inappropriate queries.
Risk Status	Identified

Risk ID	2
Risk Description	Spelling mistakes in user queries.
Category	Software.
Source	This was identified during testing of the system.
Probability	High
Impact	Medium
Strategy	Use of python package functionality which will handle this risk.
Risk Status	Identified

Risk ID	3
Risk Description	Data is not available.
Category	Software
Source	Fault in database.
Probability	Low
Impact	High
Strategy	Train the database as much possible as.
Risk Status	Identified

Risk ID	4
Risk Description	Application crash
Category	Requirements
Source	Software Design Specification documentation review.
Probability	Low
Impact	High
Response	Mitigate
Strategy	Better testing will resolve this issue.
Risk Status	Identified

- **Problem Definition and Scope:**

In order to make students efforts less, the system is developed which will give reply on behalf of human. To save the students time of coming to college or calling to the institute the system will help student to solve their issues.

- **Risk Identification:**

At the time of project execution, the system may not understand the users question or user may type wrong spelling of a word which will be not known to the system.

- **Software Requirement Specification :**

It included Softwares like Eclipse IDE, Various packages . Hardware Requirements: We do not require any hardware besides computer for programming.

### 5.3.2 Task Network

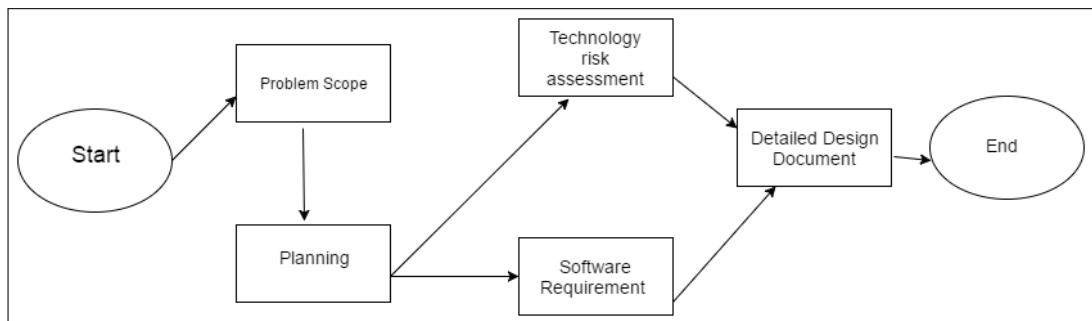


Figure 5.1: Task Network

### 5.3.3 Timeline Chart

Please refer Annex B for the timeline chart.

## 5.4 TEAM ORGANIZATION

### 5.4.1 Team structure

The team structure for the project is identified. Roles are defined.

Sr. No.	Team Members	Responsibilities
1	Aishwarya Hajare	Designing and development.
2	Priyanka Bhosale	Requirement gathering and development.
3	Guruswami Hiremath	Database management and development.
4	Rasika Nanaware	Testing and development.

Table 5.6: Team Responsibilities

#### **5.4.2 Management reporting and communication**

Regular visit to Project internal Guide to discuss various issues and finding solutions to them. Formal Technical Reviews(FTR) were conducted to discuss the overall development and future issues of the project.

**CHAPTER 6**

**SOFTWARE REQUIREMENT**

**SPECIFICATION**

## **6.1 INTRODUCTION**

### **6.1.1 Purpose and Scope of Document**

The purpose of the document is to collect and analyze all assorted ideas that have come up to define the system, its requirements with respect to consumers. Also, we shall predict and sort out how we hope this product will be used in order to gain a better understanding of the project, outline concepts that may be developed later, and document ideas that are being considered, but may be discarded as the product develops. The purpose of our product is to implement a system for recognizing road condition using arduino which will send data to the smartphone using bluetooth.

Scope : The document describes defines how our client, team and audience see the product and its functionality. It also includes non-functional requirements, design constraints and flow of data in the product. Scope of product includes : The main aim is to mark the position of road with bad conditions on the google maps.

### **6.1.2 Overview of responsibilities of Developer**

Responsibilities of Developer are to develop system which is reliable, accurate, efficient, cost effective and usable. The activities performed by our team members are:

- The requirements were gathered, analyzed, elicited, verified, and validated.
- Formal Technical Reviews or Meetings were carried to design the system and System Implementation Plan was made.
- Software Requirements Specification was provided which included user interface, hardware and software requirements, design constraints, Usage scenario of product, Flow of Activities or Operations to visualize dynamic nature of the system.
- Analysis of Functional Models and Data Flow Models were described for intended audience and client to give a view of the products functionality.

## **6.2 USAGE SCENARIO**

In our project only one User category, the simple user just enter query into system. Provide the query is an input give the specified result and give timestamp required for execution of the query.

### **6.2.1 Use-cases**

All use-cases for the software are presented:

Sr No.	Use Case	Description	Actors	Assumptions
1	Enter a query	After popping up chat bot, user will enter a query	User	The query entered by user should be proper.
2	Extract a keyword	After entering a query, keyword extraction algorithm will extract the proper keywords.	Chat Bot	Proper keyword should extract.
3	Search in database.	According to the proper keywords, the data is searched in database.	Chat Bot	Proper data should be searched in database.
4	Send reply to the user.	After successful search of data, it should be displayed to the user.	User, Chat Bot	Data is properly displayed to the user.
5	Close session	Once user closed the browser, close the current session.	User, Chat Bot	The session is closed properly.

Table 6.1: Use Cases

### 6.2.2 Use Case View

Use Case Diagram.

## 6.3 DATA MODEL AND DESCRIPTION

### 6.3.1 Data Description

Analog data is passed from accelerometer to Arduino. Serially data is passed from bluetooth module to smartphone so that it can work on it.

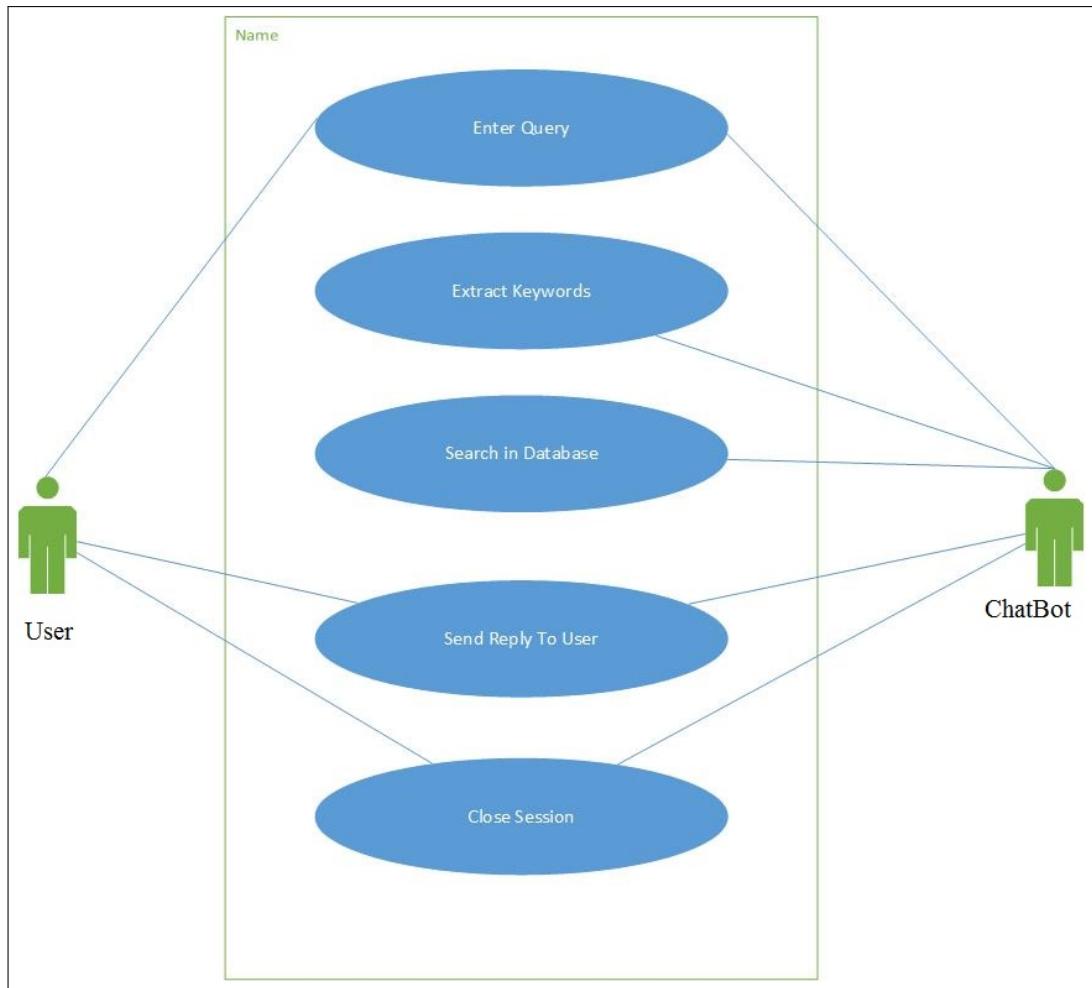


Figure 6.1: Use case diagram

### 6.3.2 Data objects and Relationships

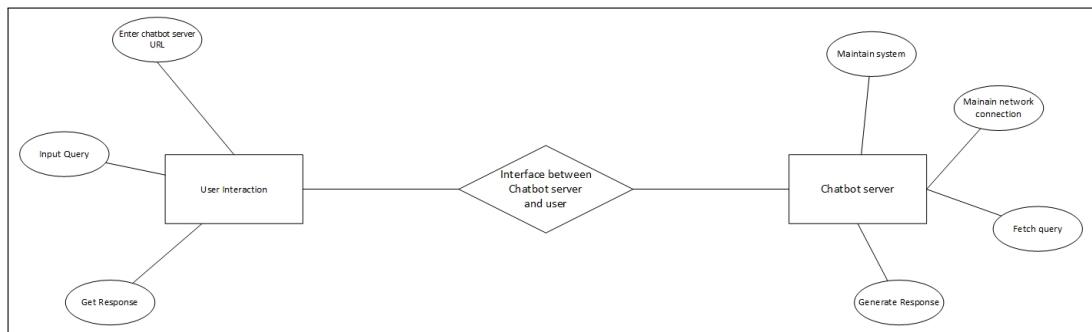


Figure 6.2: Entity Relationship Diagram

## 6.4 FUNCTIONAL MODEL AND DESCRIPTION

We will be having class **ChatBot** for generating response, getting response, setting response , class **Preprocessing** for cleaning the data, if data is not correct then convert it into appropriate format, class **Processing** for finding the correct data, finding the confidence value. **Storage** it is used for storing the data, the class **Training** for getting a processed data, and exporting the data to database, class **Application** for taking the input from user and displaying the data to the user.

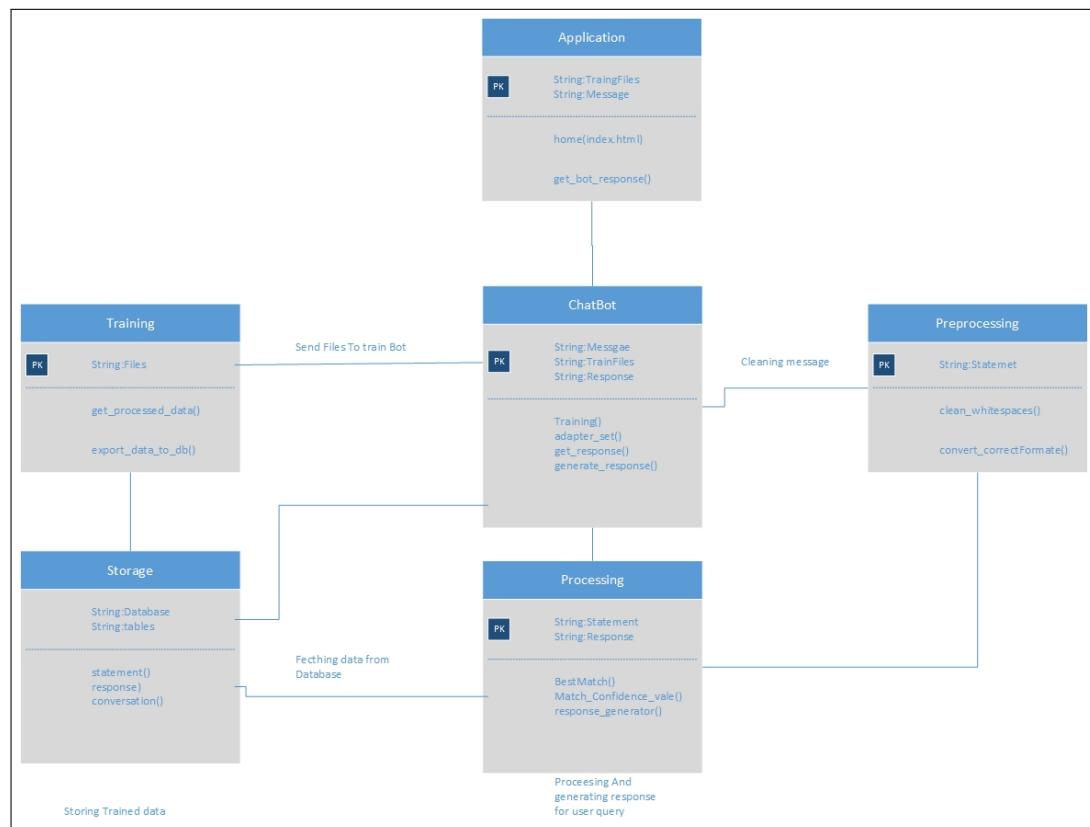


Figure 6.3: Class Diagram

### 6.4.1 Data Flow Diagram

#### 6.4.1.1 Level 0 Data Flow Diagram

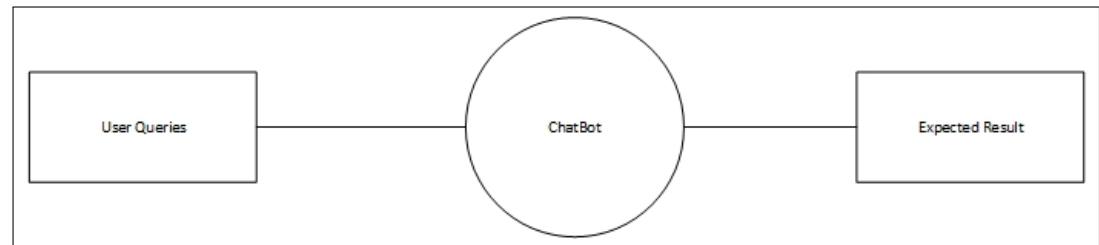


Figure 6.4: Data Flow Diagram 0

#### 6.4.1.2 Level 1 Data Flow Diagram

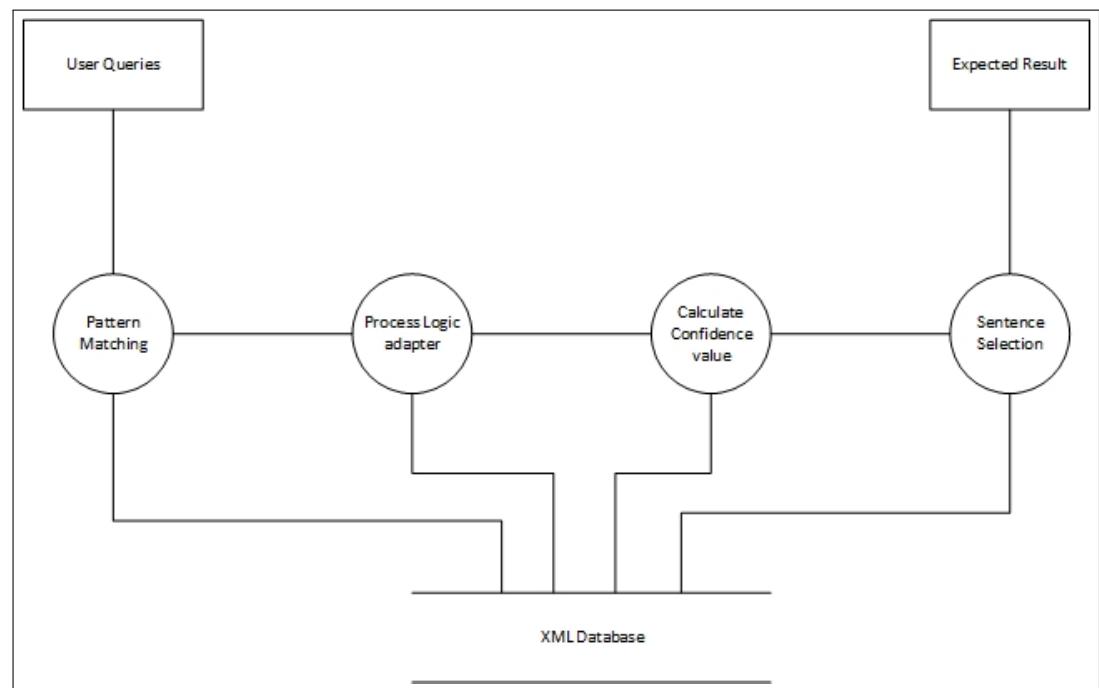


Figure 6.5: Data Flow Diagram 1

#### 6.4.2 Activity Diagram:

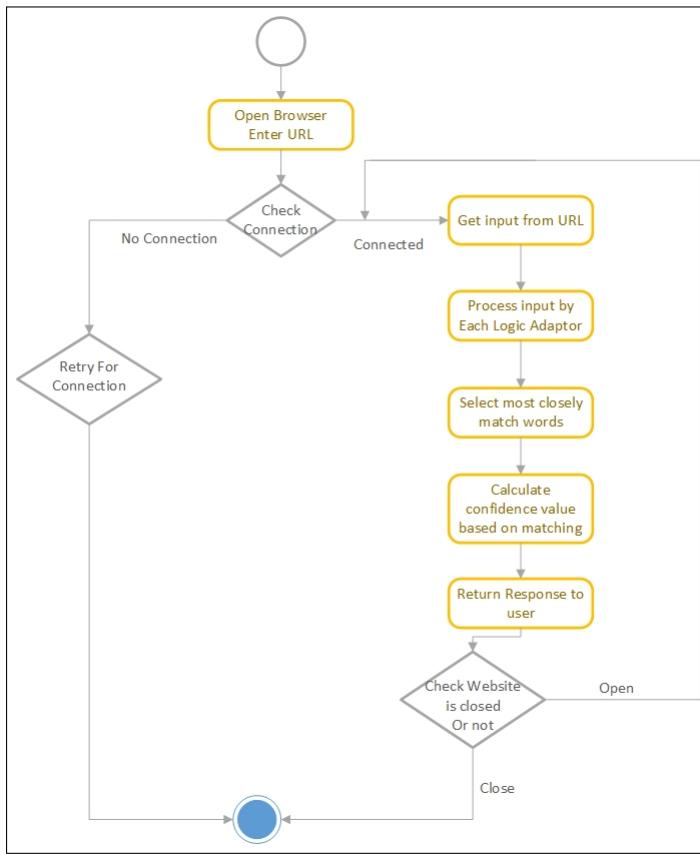


Figure 6.6: Activity diagram

#### 6.4.3 Non Functional Requirements:

- **Interface Requirement:** For perceiving input from user, UI is important which should contain a easy to understand and appealing to the customer.
- **Software quality attributes:**
  - **Availability:**  
System should be available 24\*7 to user.
  - **Modifiability:**  
User can ask any question.
  - **Performance:**  
All the queries should be answered properly.

### - Security

There is no any need for security.

#### 6.4.4 State Diagram:

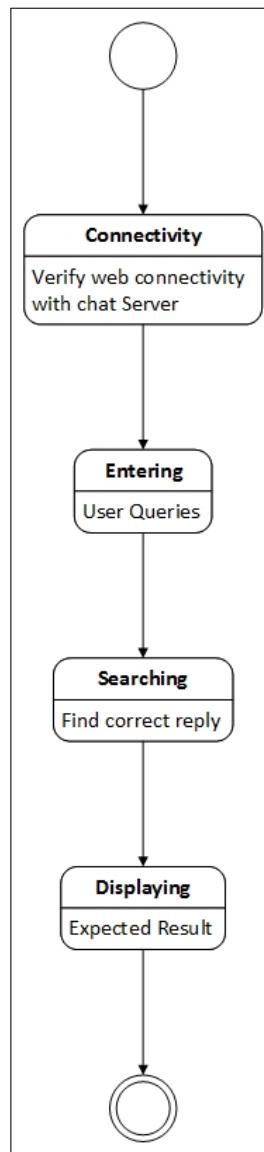


Figure 6.7: State Transition Diagram

#### 6.4.5 Design Constraints

- **Security:** Only authorized users should be having access to database.
- **Quality:** Software should have better UI and performance.

- **Time:** Project should get completed in given time span.

#### **6.4.6 Software Interface Description**

- We require the interfacing between the between the python flask framework and database. The internet connectivity should be constantly on.

# **CHAPTER 7**

## **DETAILED DESIGN DOCUMENT**

## 7.1 INTRODUCTION

This document specifies the design that is used to solve the problem of Product.

## 7.2 ARCHITECTURAL DESIGN

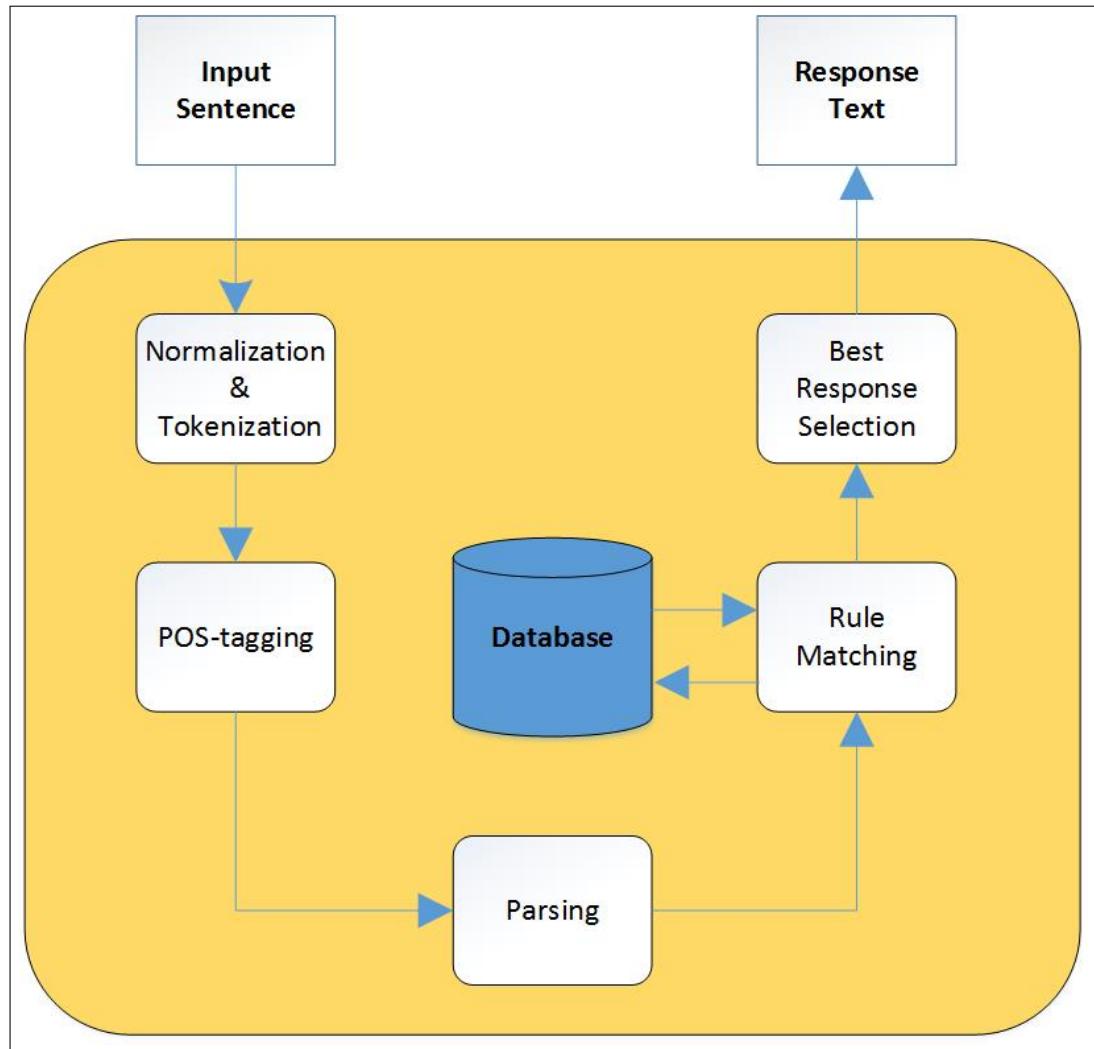


Figure 7.1: Architecture Diagram

In the architecture of our system, firstly user will enter the query. After entering the query system will tokenize the proper keywords. After finding the proper keywords, it may possible some words may have multiple meaning for e.g. charge as a noun and charge as a verb. Knowing a part of speech can help to disambiguate the meaning. The rule matching part will used for calculating the confidence value

for a given sentence. After finding the multiple values, selection of the best value is needed. So this work is done by best response selection part. And finally result should be displayed to the user.

### 7.3 COMPOENT DESIGN

#### 7.3.1 Class Diagram

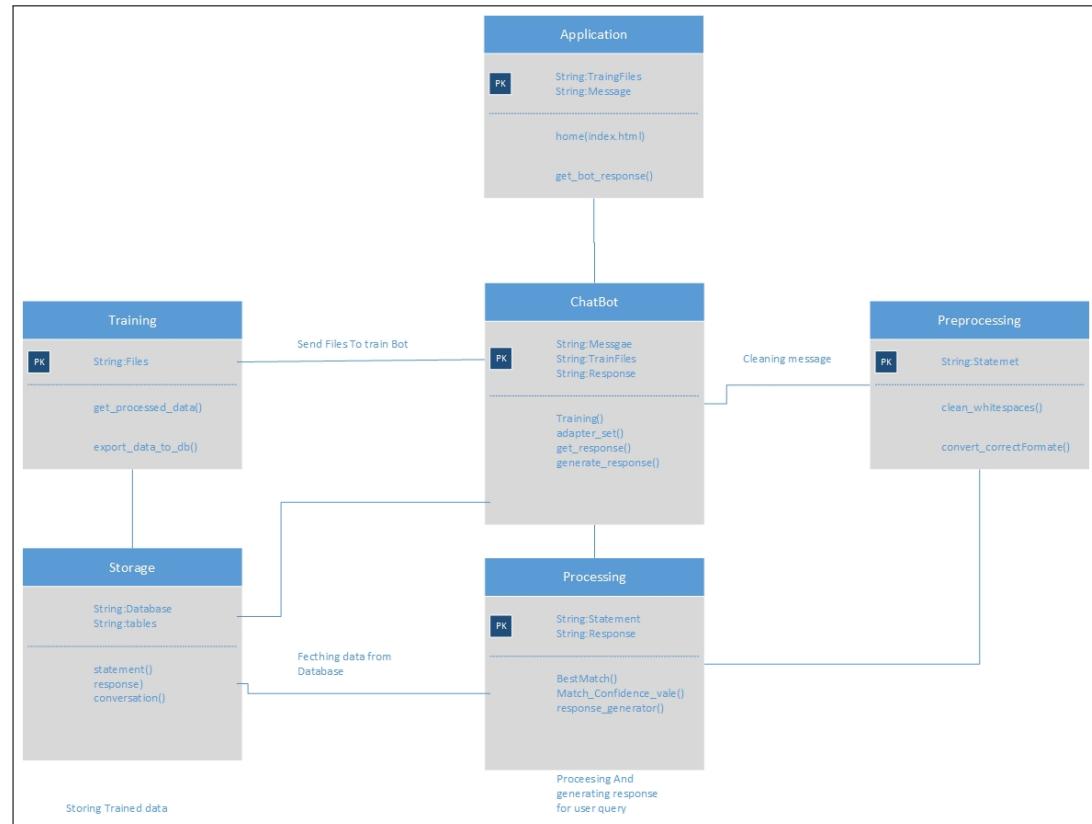


Figure 7.2: Class Diagram

We will be having class **ChatBot** for generating response, getting response, setting response , class **Preprocessing** for cleaning the data, if data is not correct then convert it into appropriate format, class **Processing** for finding the correct data, finding the confidence value. **Storage** it is used for storing the data, the class **Training** for getting a processed data, and exporting the data to database, class **Application** for taking the input from user and displaying the data to the user.

# **CHAPTER 8**

## **PROJECT IMPLEMENTATION**

## **8.1 INTRODUCTION**

We have used different types of tools and technology in our project which are given below.

## **8.2 TOOLS AND TECHNOLOGIES USED**

### **Eclipse IDE:**

Eclipse is an integrated development environment (IDE) used in computer programming, and is the most widely used Java IDE.[6] It contains a base workspace and an extensible plug-in system for customizing the environment. Eclipse is written mostly in Java and its primary use is for developing Java applications, but it may also be used to develop applications in other programming languages via plug-ins, including Ada, ABAP, C, C++, C, COBOL, D, Fortran, Haskell, JavaScript, Julia, Lasso, Lua, NATURAL, Perl, PHP, Prolog, Python, R, Ruby (including Ruby on Rails framework), Rust, Scala, Clojure, Groovy, Scheme, and Erlang. It can also be used to develop documents with LaTeX (via a TeXlipse plug-in) and packages for the software Mathematica. Development environments include the Eclipse Java development tools (JDT) for Java and Scala, Eclipse CDT for C/C++, and Eclipse PDT for PHP, among others.

### **Flask:**

Flask has many configuration values, with sensible defaults, and a few conventions when getting started. By convention, templates and static files are stored in subdirectories within the applications Python source tree, with the names templates and static respectively. While this can be changed, you usually dont have to, especially when getting started. Once you have Flask up and running, youll find a variety of extensions available in the community to integrate your project for production. The Flask core team reviews extensions and ensures approved extensions do not break with future releases.

As your codebase grows, you are free to make the design decisions appropriate for your project. Flask will continue to provide a very simple glue layer to the best

that Python has to offer. You can implement advanced patterns in SQLAlchemy or another database tool, introduce non-relational data persistence as appropriate, and take advantage of framework-agnostic tools built for WSGI, the Python web interface.

Flask includes many hooks to customize its behavior. Should you need more customization, the Flask class is built for subclassing.

### 8.2.1 Heroku :

## 8.3 METHODOLOGIES/ALGORITHM DETAILS

### 8.3.1 Algorithm

1) **Levenshtein Distance:** Levenshtein distance (LD) is a measure of the similarity between two strings, which we will refer to as the source string ( $s$ ) and the target string ( $t$ ). The distance is the number of deletions, insertions, or substitutions required to transform  $s$  into  $t$ . For example, If  $s$  is "test" and  $t$  is "test", then  $LD(s,t) = 0$ , because no transformations are needed. The strings are already identical.

If  $s$  is "test" and  $t$  is "tent", then  $LD(s,t) = 1$ , because one substitution (change "s" to "n") is sufficient to transform  $s$  into  $t$ . The greater the Levenshtein distance, the more different the strings are. The Levenshtein (or edit) distance is more sophisticated. It's defined for strings of arbitrary length. It counts the differences between two strings, where we would count a difference not only when strings have different characters but also when one has a character whereas the other does not. The formal definition follows. For a string  $s$ , let  $s(i)$  stand for its  $i$ th character. For two characters  $a$  and  $b$ , define  $r(a, b) = 0$  if  $a = b$ . Let  $r(a, b) = 1$ , otherwise. Assume we are given two strings  $s$  and  $t$  of length  $n$  and  $m$ , respectively. We are going to fill an  $(n+1)(m+1)$  array  $d$  with integers such that the low right corner element  $d(n+1, m+1)$  will furnish the required values of the Levenshtein distance  $L(s, t)$ . The definition of entries of  $d$  is recursive. First set  $d(i, 0) = i$ ,  $i = 0, 1, \dots, n$ , and  $d(0, j) = j$ ,  $j = 0, 1, \dots, m$ .

The Levenshtein distance algorithm has been used in:

Spell checking

Speech recognition

DNA analysis

Plagiarism detection

The Levenshtein distance algorithm is:

- Set n to be the length of s.
- Set m to be the length of t.
- If n = 0, return m and exit.
- If m = 0, return n and exit.
- Construct a matrix containing 0..m rows and 0..n columns.
- Initialize the first row to 0..n - Initialize the first column to 0..m.
- Examine each character of s (i from 1 to n).
- Examine each character of t (j from 1 to m).
- If  $s[i]$  equals  $t[j]$ , the cost is 0.
- If  $s[i]$  doesn't equal  $t[j]$ , the cost is 1.
- Set cell  $d[i,j]$  of the matrix equal to the minimum of:
  - a. The cell immediately above plus 1:  $d[i-1,j] + 1$ .
  - b. The cell immediately to the left plus 1:  $d[i,j-1] + 1$ .
  - c. The cell diagonally above and to the left plus the cost:  $d[i-1,j-1] + \text{cost}$ .
- After the iteration steps (3, 4, 5, 6) are complete, the distance is found in cell  $d[n,m]$ .

# **CHAPTER 9**

## **SOFTWARE TESTING**

## **9.1 TYPE OF TESTING USED**

Unit, Integration, Performance, Black-Box, white-Box, GUI, System, Recovery, Security

.

## **9.2 TESTING**

In computer hardware and software development, testing is used at key checkpoints in the overall process to determine whether objectives are being met. For example, in software development, product objectives are sometimes tested by product user representatives. When the design is complete, coding follows and the finished code is then tested at the unit or module level by each programmer; at the component level by the group of programmers involved; and at the system level when all components are combined together. At early or late stages, a product or service may also be tested for usability.

### **UNIT TESTING**

Unit testing focuses on verification effort on smallest unit of software model. Important control paths are tested to uncover the errors within the boundary of module.

The unit testing is white box oriented and can be conducted in parallel for multiple modules.

Local data structures are checked to ensure the data stored temporarily maintains its Integrity during all the steps of algorithm execution. A unit test is the lowest level of testing.

### **BLACK BOX TESTING**

Black Box testing focuses on the functional requirements of the software .It enables the software Engineer to derive sets of input conditions that will fully exercise all functional requirements for A program.

Using black box approach, a tester considers the software under test to be an opaque

box. There is no knowledge of its inner structure. The tester only has the knowledge of what it does.

The size of the software under test using this approach can vary from a simple module, member function, or subsystem to a complete system.

Several methods of black box approach are:

- Requirements based testing
- Positive and negative testing
- Boundary value analysis
- Equivalence of class partitioning
- State based testing
- Compatibility testing
- User documentation testing
- Domain testing

### **INTEGRATION TESTING:**

Once the entire module have undergone unit testing they are integrated and again tested for using a technique called integration testing.

There are two methods:

- 1.Incremental
- 2.Non incremental integration -In non incremental integration the tendency is to construct the program using Big Bang Approach.

Two different incremental strategies are:Top down and bottom up using integration.

### **Top down Integration:**

Top down integration is incremental approach. Module/component are integrated by moving downwards through control hierarchy, beginning with main control module

.

### **Bottom up Integration:**

Bottom up integration testing begin construction and testing with atomic module. Because modules are integrated from bottom up processing required for modules subordinator to given level is always available and the need for stubs is eliminated.

The steps of bottom up strategy

- Low level module are combined into clusters that perform a specific software sub functions.
- A driver is written to coordinate test cse input and output.
- The cluster is tested.
- Validation Testing
- At the culmination of integration testing software is assembled

## **PERFORMANCE TESTING:**

Performance testing is a type of testing intended to determine the responsiveness, throughput, reliability, and/or scalability of a system under a given workload. Performance testing is commonly conducted to accomplish the following: Assess production readiness. Evaluate against performance criteria.

## **SYSTEM TESTING:**

System testing of software or hardware is testing conducted on a complete, integrated system to evaluate the system's compliance with its specified requirements. System

testing falls within the scope of black-box testing, and as such, should require no knowledge of the inner design of the code or logic.

### **RECOVERY TESTING:**

Recovery testing is a type of non-functional testing technique performed in order to determine how quickly the system can recover after it has gone through system crash or hardware failure. Recovery testing is the forced failure of the software to verify if the recovery is successful.

### **SECURITY TESTING:**

Security testing is a testing technique to determine if an information system protects data and maintains functionality as intended. It also aims at verifying 6 basic principles as listed below: Confidentiality Integrity.

### **STRESS TESTING:**

Stress testing (sometimes called torture testing) is a form of deliberately intense or thorough testing used to determine the stability of a given system or entity. It involves testing beyond normal operational capacity, often to a breaking point, in order to observe the results.

Reasons can include:

1. To determine breaking points or safe usage limits
2. To confirm intended specifications are being met
3. To determine modes of failure (how exactly a system fails)
4. To test stable operation of a part or system outside standard usage

### 9.3 TEST CASES AND TEST RESULTS

Test case no	Module	Test Objective/Functionality	Normal Execution Steps	Test Data	Expected Result	Actual Result	Execution Status (Pass/Fail)
TC-01	Chat Panel	Verify that whether system is able to connect to the database	Give proper path to database connectivity	Path to database	database connectivity should be proper	Connection to the database is successful	Pass
TC-02	Chat Panel	Verify that after clicking on send button it should give proper reply	Click on send button	send button	Reply should be given	Reply given successfully	Pass

Table 9.1: Test Cases

<b>Test case no</b>	<b>Module</b>	<b>Test Objective/Functionality</b>	<b>Normal Execution Steps</b>	<b>Test Data</b>	<b>Expected Result</b>	<b>Actual Result</b>	<b>Execution Status (Pass/Fail)</b>
TC-03	Chat Panel	Verify that when user gives wrong input it should either give default value or should give correct reply	Give some wrong input in text box	Wrong sentences	Default value or correct reply gets display	same as expected.	Pass
TC-04	YML file	Verify that there is related data stored in YML file.	Data is stored	Data	Data should be stored in YML file	same as expected.	Pass
TC-05	Chat panel	Verify that the pop up and pop down functionality is working	Click on chat panel	Clicked on chat panel	the chat panel should be popped up and down	same as expected.	Pass

Table 9.2: Test Cases

# **CHAPTER 11**

## **DEPLOYMENT AND MAINTENANCE**

### **11.1 INSTALLATIONS:**

1. Open any web browser
2. Type AISSMS Chatbot web address in search bar.
3. Click on the site.
4. you are ready with the chat bot conversation.

## **CHAPTER 12**

## **CONCLUSION AND FUTURE SCOPE**

## **12.1 CONCLUSION:**

Chat Bot is a simple and cost effective system to give automatic reply to the students queries. Because of this user doesn't need to visit the college or call the institute for enquiry. It will save the time of students. Searching a single information from large data is quite hard task for students. So this functionality also saves the time of the student. Student will enter a query in a chat panel which will be taken further for next processing. It will extract the keywords from entered query. Based on that keyword it will search the data from database. And it will produce the output. The future work would be online fetching of the data.

## **12.2 FUTURE SCOPE:**

Our system is limited the data till static data base only. The data which is not available in our local data base should be fetched from online sources. This will increase the response generation quality of the system.

## **CHAPTER 13**

## **REFERENCES**

- K. J. Oh, D. Lee, B. Ko and H. J. Choi, "A Chatbot for Psychiatric Counseling in Mental Health Care Service Based on Emotional Dialogue Analysis and Sentence Generation," 2017 18th IEEE International Conference on Mobile Data Management (MDM), Daejeon, 2017, pp. 371-375
- Prof. Nikita Hatwar chat bot for marketing field named as AI based chat bot.In 2016
- Y. Wu, G. Wang, W. Li and Z. Li, "Automatic Chatbot Knowledge Acquisition from Online Forum via Rough Set and Ensemble Learning," 2008 IFIP International Conference on Network and Parallel Computing, Shanghai, 2008, pp. 242-246.
- M. N. Kumar, P. C. L. Chandar, A. V. Prasad and K. Sumangali, "Android based educational Chatbot for visually impaired people," 2016 IEEE International Conference on Computational Intelligence and Computing Research (ICCIC), Chennai, 2016, pp. 1-4.
- J. C. De Oliveira, D. H. Santos and M. P. Neto, "Chatting with Arduino platform through Telegram Bot," 2016 IEEE International Symposium on Consumer Electronics (ISCE), Sao Paulo, 2016, pp. 131-132
- Randall Wald, Taghi M. Khoshgoftaar, Amri Napolitano, Chris Sumner, "Predicting susceptibility to social bots on Twitter", Information Reuse and Integration (IRI) 2013 IEEE 14th International Conference on, pp. 6-13, 2013.
- W. Zhu and C. Lee, "Internet security protection for IRC-based botnet," 2015 IEEE 5th International Conference on Electronic Information and Emergency Communication, Beijing, 2015, pp. 63-66.
- Qiuze Yu, Shunxin Min, Bo Pang and Yan Zhang, "Simultaneously Multiple-object Pattern Matching based on Multi-swarms Particle Swarm Optimization," 2017 2nd International Conference on Image, Vision and Computing (ICIVC), Chengdu, 2017, pp. 122-127

- M. U. Nisar, S. Voghoei and L. Ramaswamy, "Caching for Pattern Matching Queries in Time Evolving Graphs: Challenges and Approaches," 2017 IEEE 37th International Conference on Distributed Computing Systems (ICDCS), Atlanta, GA, 2017, pp. 2352-2357.
- D. G. Bailey, "Identification of Trax threats using pattern matching," 2016 International Conference on Field-Programmable Technology (FPT), Xi'an, 2016, pp. 319-322.

## **ANNEXURE A**

# **LABORATORY ASSIGNMENTS ON PROJECT ANALYSIS OF ALGORITHMIC DESIGN**

- Idea matrix is a strategy to represent the basic ideas of project. In IDEA:
  - I[Strategic]
    - 1.Increase:Tells which aspects need to be increased for building knowledge and understanding the system.
    - 2.Improve:The skills needed to be innovated.
    - 3.Ignore:Parameters which can be and that do not affect the results.
  - D[Functional]
    - 1.Drive:Factors that help in building the system.
    - 2.Deliver:The factors delivered at different levels.
    - 3.Decrease:Factors decreasing exponentially.
  - E[knowledge building]
    - 1.Educate:Educate to improve knowledge efficiency.
    - 2.Evaluate:Parameters that are needed to be evaluated.
    - 3.Eliminate:Elements that should be eliminated to overcome the hindrance in knowledge building.
  - A
    - 1.Accelerate:Elements that need to be accelerated beyond regular speed.
    - 2.Associate:Factors that are needed to be associated to be knowledge build.
    - 3.Avoid:Elements which are to be avoided.

I	D	E	A
Increase:Accuracy of Algorithm	Drive:Improved algorithm	Educate:End user	Accelerate:Internet access
Improve:Knowledge about Python	Deliver:Appropriate response	Evaluate: Response	Associate:Value creation, time saving, power source
Ignore:Inaccurate values	Decrease:Irrelevant response	Eliminate:Inaccurate response	Avoid:Initial cost, competition-centric approach, mere value-driven approach

Table A.1: IDEA Matrix

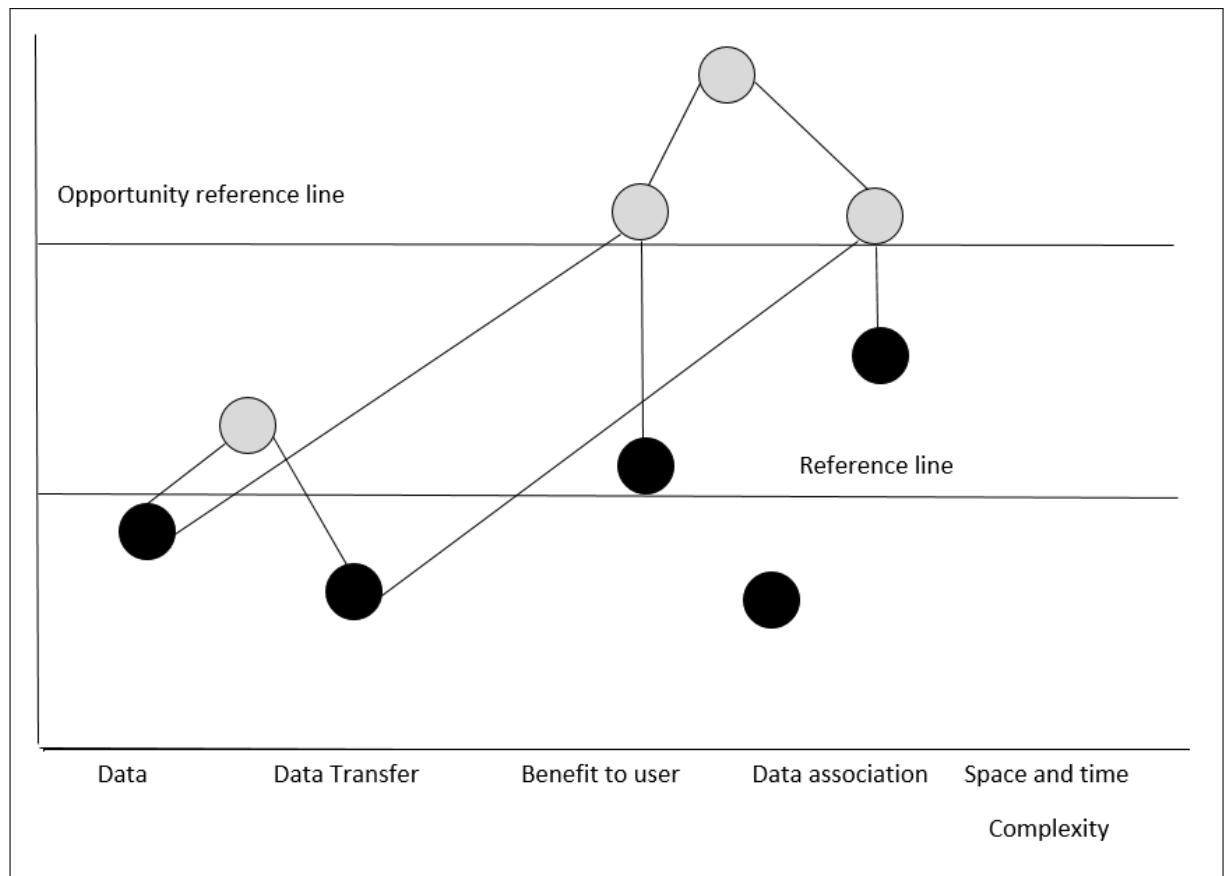


Figure A.1: Canvas Graph

Test Case No.	Expected Input	Expected Output	Input Example
T1	Reading the data from students given query	Text is displayed on screen	how are you?
T2	Reading the data from students given query	Text is displayed on screen	hello
T3	Any random input	System should stop processing and display error message	hw r u?

Table A.2: Test Cases for Black Box testing

**ANNEXURE B**

**PROJECT PLANNER**

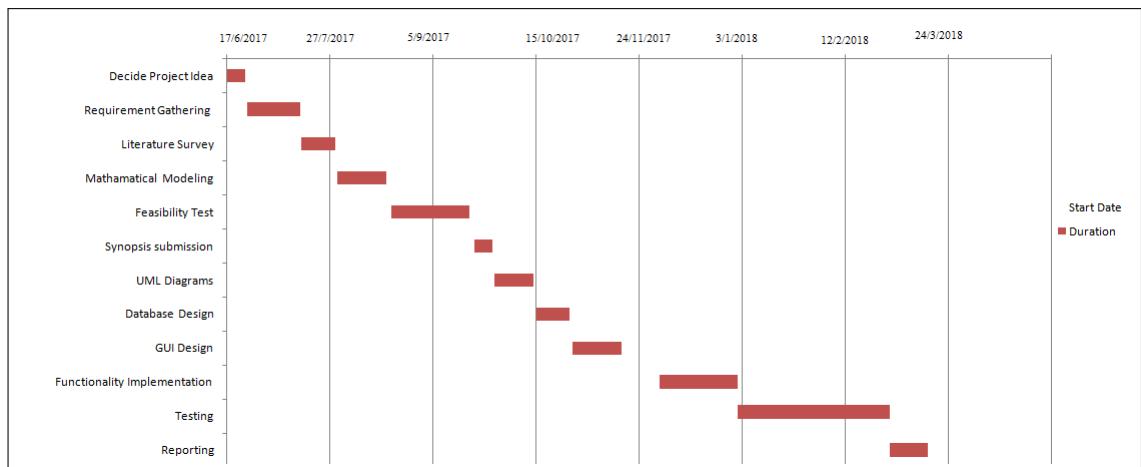


Figure B.1: Project Plan

**ANNEXURE C**

**REVIEWERS COMMENTS OF PAPER**

**SUBMITTED**

## **C.1 REVIEW OF CONFERENCE/JOURNAL PAPERS SUPPORTING PROJECT IDEA**

1. Paper Title:Chat Bot for education system.

Name of the Conference/Journal where paper submitted :International Journal Of Emerging Technology and Computer Science  
Paper accepted/rejected : Accepted.

2. Paper Title:Chat Bot for education system.

Name of the Conference/Journal where paper submitted :International Journal Of Advance Engineeringand Research Development(IJAERD)  
Paper accepted/rejected : Accepted.

**ANNEXURE D**

**PLAGIARISM REPORT**

## Chat Bot For Education System

---

### ORIGINALITY REPORT

---

**17** % SIMILARITY INDEX    31% INTERNET SOURCES    15% PUBLICATIONS    % STUDENT PAPERS

---

### PRIMARY SOURCES

---

<b>1</b>	<a href="http://en.wikipedia.org">en.wikipedia.org</a> Internet Source	1 %
<b>2</b>	<a href="http://www.garybeene.com">www.garybeene.com</a> Internet Source	1 %
<b>3</b>	<a href="http://keshavatech.com">keshavatech.com</a> Internet Source	1 %
<b>4</b>	<a href="http://docplayer.net">docplayer.net</a> Internet Source	1 %
<b>5</b>	<a href="http://www.cut-the-knot.org">www.cut-the-knot.org</a> Internet Source	1 %
<b>6</b>	<a href="http://www.exforsys.com">www.exforsys.com</a> Internet Source	1 %
<b>7</b>	<a href="http://www.authorstream.com">www.authorstream.com</a> Internet Source	1 %
<b>8</b>	<a href="http://blablawriting.com">blablawriting.com</a> Internet Source	1 %
<b>9</b>	<a href="http://www.automationlearn.com">www.automationlearn.com</a> Internet Source	1 %

---

<b>10</b>	Saudagar L. Jadhav, Manisha P. Mali. "Pre-Recommendation Clustering and Review Based Approach for Collaborative Filtering Based Movie Recommendation", International Journal of Information Technology and Computer Science, 2016 Publication	1 %
<b>11</b>	<a href="http://www.cc.gatech.edu">www.cc.gatech.edu</a> Internet Source	1 %
<b>12</b>	Zhu, Wenzheng, and Changhoon Lee. "Internet security protection for IRC-based botnet", 2015 IEEE 5th International Conference on Electronics Information and Emergency Communication, 2015. Publication	1 %
<b>13</b>	<a href="http://www.technotrice.com">www.technotrice.com</a> Internet Source	1 %
<b>14</b>	<a href="http://www.badtesting.com">www.badtesting.com</a> Internet Source	1 %
<b>15</b>	<a href="http://toc.proceedings.com">toc.proceedings.com</a> Internet Source	1 %
<b>16</b>	<a href="http://flask.pocoo.org">flask.pocoo.org</a> Internet Source	1 %
<b>17</b>	<a href="http://en.m.wikipedia.org">en.m.wikipedia.org</a> Internet Source	1 %

---

**ANNEXURE E**

**TERM-II PROJECT LABORATORY**

**ASSIGNMENTS**

1. At the very first we developed a GUI with a login logout facility. But latter on we decided to depricate that functionality. We thought it is not our main processing. The main processing is extracting the keyword, searching the data and displaying the response. In stead of main processing we focused on core functionalities.
2. We have also tested our project using different testing methods.