SAVITRIBAI PHULE PUNE UNIVERSITY

A PROJECT REPORT ON

Emotionally Intelligent Chatbot

SUBMITTED TOWARDS THE
FULFILLMENT OF THE REQUIREMENTS OF

BACHELOR OF ENGINEERING
(Computer Engineering)

BY

SAKSHI DUBEWAR	71941846B
SHREEYASH KHALATE	71809414L
SHUBHAM KOKANE	71809420E
SHREYA NEWALE	71941854C

Under The Guidance of

Prof. Jameer Kotwal

DEPARTMENT OF COMPUTER ENGINEERING

Pimpri Chinchwad College of Engineering and Research

Ravet,Pune-412101.

Pimpri Chinchwad College Of Engineering & Research
DEPARTMENT OF COMPUTER ENGINEERING

CERTIFICATE

This is to certify that the Project Entitled

Emotionally Intelligent Chatbot

Submitted by

SAKSHI DUBEWAR	71941846B
SHREEYASH KHALATE	71809414L
SHUBHAM KOKANE	71809420E
SHREYA NEWALE	71728404C

have successfully completed the project entitled ***"Emotionally Intelligent Chatbot "*** in the fulfillment of B. E. (Computer Engineering) and this work has been carried out in presence of Prof. Jameer Kotwal.

Prof. Jameer Kotwal
Internal Guide

Prof. Jameer Kotwal
Project Co-ordinator

Prof. Dr. Archana Chaugule
H.O.D

Prof. Dr. Harish Tiwari
Principal

External Examiner

PROJECT APPROVAL SHEET

A Project Title

Emotionally Intelligent Chatbot

Is successfully completed by

SAKSHI DUBEWAR	71941846B
SHREEYASH KHALATE	71809414L
SHUBHAM KOKANE	71809420E
SHREYA NEWALE	71728404C

at

DEPARTMENT OF COMPUTER ENGINEERING

(Pimpri Chinchwad College Of Engineering And Research)

SAVITRIBAI PHULE PUNE UNIVERSITY,PUNE

ACADEMIC YEAR 2020-2021

Prof. Jameer Kotwal
Internal Guide
Dept. of Computer Engg.

Prof. Archana Chaugule
H.O.D
Dept. of Computer Engg.

Abstract

An emotionally intelligent chatbot system aims to make an effective conversation between humans and machine in as natural and interactive manner as possible. The chatbot agent has pre-embedded knowledge base to identify the sentences, intents, entities and context of the input query to be precise for making a valid, predictable decision itself as a self-generated response to answer the query.The present technical project consist of developing an intelligent system for college enquiry purposes using a web-based chatbot agent, through Machine Learning and query processing and virtual emotion recognition system to analyse the sentiment of the visitor towards this interaction and his satisfaction for the service.

Emotionally Intelligent College Enquiry Chatbot System is nothing but chatbot to understand the user queries and respond to it during a conversation. Chatbot can actively helps human to involve in a digital automated conversation with a machine or a system with effective Speech Emotion Recognition (SER), which can be stated as extraction of emotional state of the speaker from his/her text based or voice based speech signal through his/her input query. In this the processing our system recognizes the user's emotional state using computational analysis of speech features. It is an act of attempting to recognize human emotion and effective states from the input query.

Keywords: Emotion Classification, Chatbot Agent, Sentiment Analysis, Multinomial Naive Bayes, NLTK.

Acknowledgments

It gives us great pleasure in presenting the project report on ***'Emotionally Intelligent Chat-bot'****.*

I would like to take this opportunity to thank my internal guide ***Prof. Jameer Kotwal*** *for giving me all the help and guidance I needed. I am really grateful to them for their kind support. Their valuable suggestions were very helpful.*

I am also grateful to ***Prof. Archana Chaugule****, Head of Computer Engineering Department, Pimpri Chinchwad College Of Engineering and Research for hes indispensable support, suggestions.*

In the end our special thanks to ***to all staff members*** *for providing various resources such as laboratory with all needed software platforms, continuous Internet connection, for Our Project.*

Sakshi Dubewar
Shreeyash Khalate
Shubham Kokane
Shreya Newale
(B.E. Computer Engg.)

Contents

List of Figures

List of Tables

Chapter 1

Synopsis

1.1 PROJECT TITLE

”Emotionally Intelligent Chatbot”

1.2 PROJECT OPTION

Final Year Project

1.3 INTERNAL GUIDE

Prof. Jameeer Kotwal

1.4 SPONSORSHIP AND EXTERNAL GUIDE

No

1.5 TECHNICAL KEYWORDS (AS PER ACM KEYWORDS)

1. Emotion Classification
2. Chatbot Agent
3. Sentiment Analysis
4. Multinomial Naive Bayes
5. NLTK

1.6 PROBLEM STATEMENT

To design and implement an emotionally intelligent college enquiry chatbot system for the viewers of the college website, performing sentiment analysis and emotion classification on the conversation history of the user-bot interactions to determine his/her level of interest in the college.

1.7 ABSTRACT

An emotionally intelligent chatbot system aims to make an effective conversation between humans and machine in as natural and interactive manner as possible. The chatbot agent has pre-embedded knowledge base to identify the sentences, intents, entities and context of the input query to be precise for making a valid, predictable decision itself as a self-generated response to answer the query. Emotionally Intelligent College Enquiry Chatbot System is nothing but chatbot to understand the user queries and respond to it during a conversation. Chatbot can actively helps human to involve in a digital automated conversation with a machine or a system with effective Speech Emotion Recognition (SER), which can be stated as extraction of emotional state of the speaker from his/her text based or voice based speech signal through his/her input query.

1.8 GOALS AND OBJECTIVES

- Goal:
  - The aim is to design a college enquiry chatbot system to solve the queries of the users within the context of the respective college.
  - Another goal is to classify the emotion of the users based on their input queries in order to determine his/her interest in the college.
  - To recognize the sentiment of the user asking the queries

- Objective:

  This project aims to solve the queries of users through a chatbot interface with text/voice based input query.

  1. The user asks the query in text or voice based format.

  2. The query gets further processed through intent classifier and entity extractor and valid response is generated.

  3. Further all the queries (conversation history) are used to classify overall emotions and sentiments in users about the college.

  The formulation of research objectives eventually led to the design of a suitable chat bot agent and construct a research methodology to explore and investigate the research issues in order to achieve the stated objectives.

1.9 RELEVANT MATHEMATICS ASSOCIATED WITH THE PROJECT

- Input: User's query

- Output: Valid response for the query, Analyzed sentiment and emotion classification on overall conversation history stored in the resulting dataset

- Intent and context classification, entity extraction by Dialogflow's two simultaneous algorithms to match intents: rule-based grammar matching and ML matching

- Relevant algorithms for speech emotion recognition are Multinomial Naive Bayes, k-Nearest Neighbor (k-NN), Logistic Regression

- Functions : Insert query, Getting appropriate response

- Mathematical formulation if possible:
  S=U,A,I,O,T1,T2,Su,F
  Where, S=System
  U=Set of website/bot console viewers.
  A=Admin.
  I=Set of input queries.
  O=Output response.
  T1=Task for generating outcomes
  T2=Task for emotion classification.
  Su=Response generated
  F=Response not generated.

- Success Conditions: Queries should be correctly typed and valid response should be generated..

- Failure Conditions: Queries when incorrectly entered are not accepted and response is not generated

1.10 NAMES OF CONFERENCES / JOURNALS WHERE PAPERS CAN BE PUBLISHED

- Journal Of University Of Shanghai For Science And Technology (JUUST)
- International Journal of Recent Technology and Engineering (IJRTE)
- International Research Journal of Engineering and Technology IRJET Journal

1.11 REVIEW OF CONFERENCE/JOURNAL PAPERS SUPPORTING PROJECT IDEA

- **Ms.Ch.Lavanya Susanna, R.Pratyusha, P.Swathi, P.Rishi Krishna, V.Sai Pradeep," COLLEGE ENQUIRY CHATBOT", International Research Journal of Engineering and Technology (IRJET),on 2020**

  This paper proposes to develop an algorithm which will be wont to identify answers associated with user submitted questions. To develop a database were all the related data are going to be stored and to develop an internet interface.. The user will not waste a lot of time searching for the acceptable notices.

- **Ali Harimi, Ali Shahzadi, Alireza Ahmadyfard, Khashayar Yaghmaie, "Classification of emotional speech using spectral pattern features", Journal of Artificial Intelligence and Data Mining,on 2014**

  In this paper, we propose an algorithms for the recognition of human emotions from speech.Also, finding effective features for classifying valence related emotions and recognize emotions for real work data, evaluating the proposed system under different conditions such as the presence of noise and chatter .

- **Sumit Wailthare, Tushar Gaikwad, Ketkee Khadse, Pooja Dubey,"Artificial Intelligence Based Chat-Bot",International Research Journal of Engineering and Technology (IRJET),on 2018**

  This paper has to be provide answers to the query of the user very effectively. The objective of the project is to develop an algorithm that will be used to identify answers related to user submitted questions. The need is to develop a database where all the related data will be stored and to develop a web interface.

- **Gustavo Assuncao, Paulo Menezes, Fernando Perdigao, "Speaker Awareness for Speech Emotion Recognition", International Journal Of Online And Biomedical Engineering, 2020**

  In this paper evaluate a large scale machine learning model for classification of emotional states trained for speaker identification but is instead used here as a front-end for extracting robust features from emotional speech. We aim to verify that SER improves when some speaker's emotional prosody cues are considered.

1.12 PLAN OF PROJECT EXECUTION

The plan of project execution gives idea for the time required for project to complete.

Plan of Project Execution

Action plan outline	Goal
Requirement gathering	Gathered requirement for development of project
Designing	Designing for the project
Implementation	Coding of working process of project
Testing	Testing the project at different phases of development.
Deployment	Merging of different tested modules in a single unit for making it ready to use for end user
Maintenance and updating	Maintaining the developed project for the given time and also up- gradation

Month	Activity
July 2020	Searching And Selection of Domain And Topic And Building an abstract idea of a problem
August 2020	Literature Survey, Problem Definition
September 2020	Requirement gathering and Project planning
October 2020	Planning and deciding the platform/technology for the project
November 2020	Research Paper Analysis And Designing The Chatbot
December 2020	Testing the conversational flow of the chatbot
January 2021	Implementation of the emotion classification and Analytics part in the backend
February 2021	Testing of all modules in the project
March 2021	Research paper preparation and publication
April 2021	Final Project Evaluation
May 2021	Project Report Making
June 2021	Submission Of Project Report

Following is the PERT chart of our project.

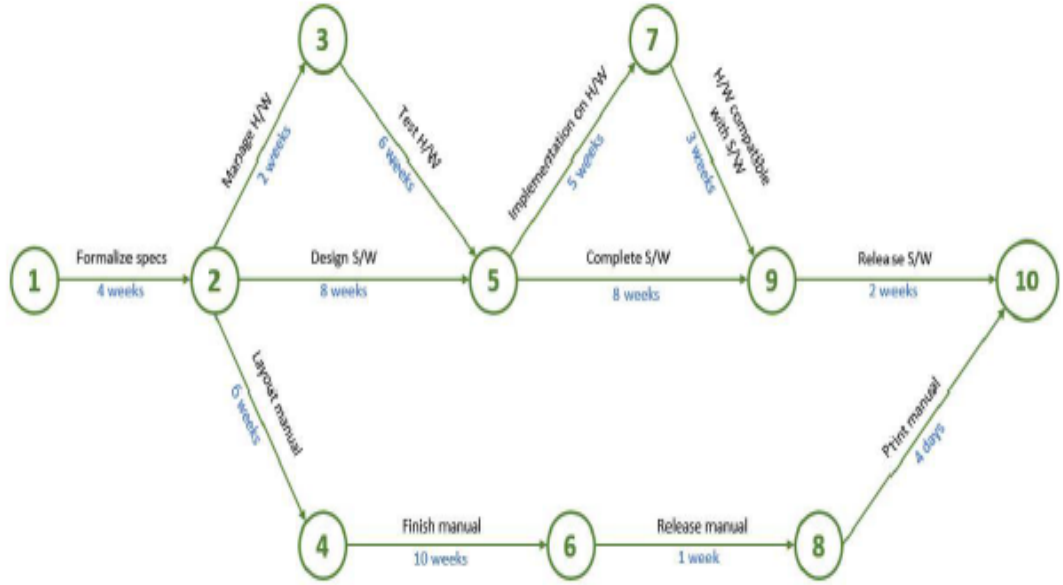


Figure 1.1: PERT chart

Chapter 2

Technical Keywords

2.1 AREA OF PROJECT

- Speech Emotion Recognition:-

  Speech Emotion Recognition (SER) can be defined as extraction of the emotional state of the speaker from his or her text based or voice based speech signal. It is the act of attempting to recognize human emotion and affective states from speech. This is capitalizing on the fact that text conversations of a user often reflects underlying emotion through certain keywords. This is also the phenomenon that animals like dogs and horses employ to be able to understand human emotion. Emotional speech processing technologies recognize the user's emotional state using computational analysis of conversational features. Our project is designing a chatbot and building a supervised machine learning model to classify emotion and sentiment on the conversation history of users.

2.2 TECHNICAL KEYWORDS

1. Emotion Classification
2. Chatbot Agent
3. Sentiment Analysis
4. Multinomial Naive Bayes
5. NLTK

Chapter 3

Introduction

3.1 PROJECT IDEA

- Textual conversational agent or chat-bots development gather tremendous traction from both academia and industries in recent years. Nowadays, chat-bots are widely used as an agent to communicate with a human in some services such as booking assistant, customer service, and also a personal partner. Emotion recognition in spoken dialogues has been gaining increasing interest all through current years.
- Speech Emotion Recognition is a hot research topic in the field of Human Computer Interaction. It has a potentially wide applications, such as the interface with robots, banking, call centers,car board systems, computer games etc. That is why Emotionally intelligent chatbot and voice based reognition came into picture in our research work.

3.2 MOTIVATION OF THE PROJECT

Emotion plays a significant role in daily interpersonal human interactions. This is essential to our rational as well as intelligent decisions. It helps us to match and understand the feelings of others by conveying our feelings and giving feedback to others. Research has revealed the powerful role that emotion play in shaping human social interaction. Emotional displays convey considerable information about the mental state of an individual. This has opened up a new research field called automatic emotion recognition, having basic goals to understand and retrieve desired emotions.

3.3 LITERATURE SURVEY

- **Ms.Ch.Lavanya Susanna, R.Pratyusha, P.Swathi, P.Rishi Krishna, V.Sai Pradeep," COLLEGE ENQUIRY CHATBOT", International Research Journal of Engineering and Technology (IRJET),on 2020**

  This paper proposes to develop an algorithm which will be wont to identify answers associated with user submitted questions. To develop a database were all the related data are going to be stored and to develop an internet interface.. The user will not waste a lot of time searching for the acceptable notices.

- **A. Harimi1*, A. Shahzadi1, A.R. Ahmadyfard2 and Kh.Yaghmaie Department of Electrical Engineering and Robotics, Shahrood University of technology, Iran 09 February 2013**

  This paper proposes to evaluate the proposed SPs and HEs for the recognition of human emotions from speech.Aim of a speech emotion recognition system is to recognize emotions for real work data, evaluating the proposed system under different conditions such as the presence of noise and chatter is useful.

- **Sumit Wailthare, Tushar Gaikwad, Ketkee Khadse, Pooja Dubey, "Artificial Intelligence Based Chat-Bot" , Dept. Of Computer Science Technology, Nagpur Institute of Technology, Nagpur, Maharashtra, India 03 (IRJET) — Mar-2018,**

  This paper has to be provide answers to the query of the user very effectively. The objective of the project is to develop an algorithm that will be used to identify answers related to user submitted questions. The need is to develop a database where all the related data will be stored and to develop a web interface.

- **Gustavo Assunção , Paulo Menezes Institute of Systems and Robotics, Coimbra, Portugal , "Speaker Awareness for Speech Emotion Recognition" (2020)**

  In this paper evaluate a large scale machine learning model for classification of emotional states trained for speaker identification but is instead used here as a front-end for extracting robust features from emotional speech. We aim to verify that SER improves when some speaker's emotional prosody cues are considered.

Chapter 4

Problem Definition and Scope

4.1 PROBLEM STATEMENT

To design and implement an emotionally intelligent college enquiry chatbot system for the viewers of the college website, performing sentiment analysis and emotion classification on the conversation history of the user-bot interactions to determine his/her level of interest in the college.

4.1.1 Goals and objectives

- Goal:
  - The aim is to design a college enquiry chatbot system to solve the queries of the users within the context of the respective college.
  - Another goal is to classify the emotion of the users based on their input queries in order to determine his/her interest in the college.
  - To recognize the sentiment of the user asking the queries

- Objective:

  This project aims to solve the queries of users through a chatbot interface with text/voice based input query.

  1. The user asks the query in text or voice based format.

  2. The query gets further processed through intent classifier and entity extractor and valid response is generated.

  3. Further all the queries (conversation history) are used to classify overall emotions and sentiments in users about the college.

  The formulation of research objectives eventually led to the design of a suitable chat bot agent and construct a research methodology to explore and investigate the research issues in order to achieve the stated objectives.

4.1.2 Statement of Scope

In future the accuracy of the proposed system can be increased by combination of feature extraction techniques such as TF-IDF and Count vectorization. The Chatbot agent can be modified to add more functional entities like booking a college appointment, online college ERP login, etc. through chatbot agent itself.

4.2 SOFTWARE CONTEXT

The proposed system will be used by the users visiting our college website to enquire about the college. The viewer need not have to visit the college campus or navigate through complete website for his/her enquiry related issues. Instead the user will get automated response from our chatbot agent for his/her query. The emotion classification and setiment analysis part will be done on the conversation history to be retreived manually through google cloud platform.

4.3 OUTCOME

The outcome will be the valid auto-generated response from the chatbot system for the user's input query.

4.4 APPLICATIONS

- This system can be used for organizational websites like colleges, schools, passport enquiry website and many more government websites too.
- This system can be applicable for various social media platforms too to generate automated responses in their conversations like Facebook Messenger, Slack, Telegram, Text Messages, etc

4.5 HARDWARE RESOURCES REQUIRED

- CPU Speed : 2GHz
- RAM : 8GB
- Hard Disk : 1TB

4.6 SOFTWARE RESOURCES REQUIREMENT

- Operating System : Windows
- Languages : HTML, CSS, Javascript, Python
- Backend : Google Firestore Database (used by Dialogflow)
- Server : Dialogflow Webhook
- Browser : Internet Explorer, Mozilla Firefox, Google Chrome

Chapter 5

Project Plan

5.1 PROJECT ESTIMATES

5.1.1 Reconciled Estimates

Cost Estimate

Not applicable

Time Estimates

Approximately 11 months

5.1.2 Project Resources

1. Communicated with the college students to know about the challenges experienced by them in the past in college enquiry process.
2. Internet Services
3. Team members availability
4. Equipment and supplies

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

This section discusses Project risks and the approach to managing them.

5.2.1 Risk Identification

The probability risks in the project are less, but nothing is perfect so risks may occur. Our project problem is NP complete. Following risks have been identified:

- Input query may contain a lot of spelling mistakes.
- System wouldn't classify intent in some cases when query is structured semantically different or grammatically incorrect.
- Inappropriate query formation - To overcome this risk, we'll define sufficient intent classes.

Following is the questionnaire:

1. Are requirements fully understood by the software engineering team ?

Yes

2. Have customers been involved fully in the definition of requirements?

No

3. Do end-users have realistic expectations?

Yes

4. Does the software engineering team have the right mix of skills?

Yes

5. Are project requirements stable?

Yes

6. Is the number of people on the project team adequate to do the job?

Yes

5.2.2 Risk Analysis

The risks for the Project can be analyzed within the constraints of time and quality

5.2.3 Overview of Risk Mitigation, Monitoring, Management

The identification of Risk is central to the success and failure of the project, hence we have made a concentrated effort to minimize and even eliminate. Certain risk related. Software risk could be classified into categories. Internal and External risk, those risk which arise

from the risk factor within the organization can be defined internal risk and the risk coming from outside is called external risk. Internal risk avoidance can be done by clear picturing the process, product risk. Following are the details for each risk.

ID	Risk Description	Probability	Impact		
			Schedule	Quality	Overall
1	Description 1	Low	Low	High	High
2	Description 2	Low	Low	High	High

Table 5.1: Risk Table

Probability	Value	Description
High	Probability of occurrence is	$> 75\%$
Medium	Probability of occurrence is	$26 - 75\%$
Low	Probability of occurrence is	$< 25\%$

Table 5.2: Risk Probability definition

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 5.3: Risk Impact definitions

Risk ID	1
Risk Description	Hardware Failure
Category	Development Environment.
Source	This was identified during early devolopment and testing.
Probability	Low
Impact	High
Response	Accept
Strategy	Reliable hardware will overcome this
Risk Status	Occurred

Table 5.4: Risk Id 1

Risk ID	2
Risk Description	Connection Failure
Category	Development Environment.
Source	This was identified during early development and testing.
Probability	Low
Impact	High
Response	Mitigate
Strategy	Proper authentication and proper Internet connectivity will resolve this issue
Risk Status	Identified

Table 5.5: Risk Id 2

5.3 PROJECT SCHEDULE

5.3.1 Project Task Set

Major Tasks in the Project stages are:

- Task 1: Developing separate modules of a basic chatbot system, emotion recognition and textual sentiment analysis system.
- Task 2: Combining the developed modules.
- Task 3: Making proper connection to database.
- Task 4: Integration with website.
- Task 5: Testing

5.3.2 Task Diagram

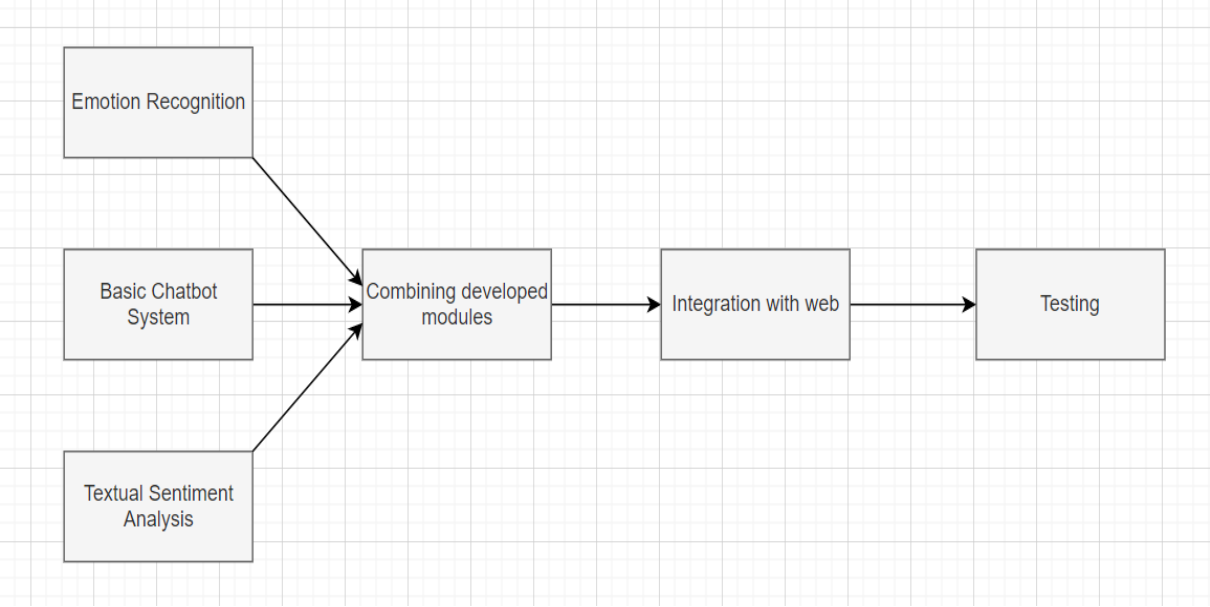


Figure 5.1: Task Network

5.3.3 Timeline Chart

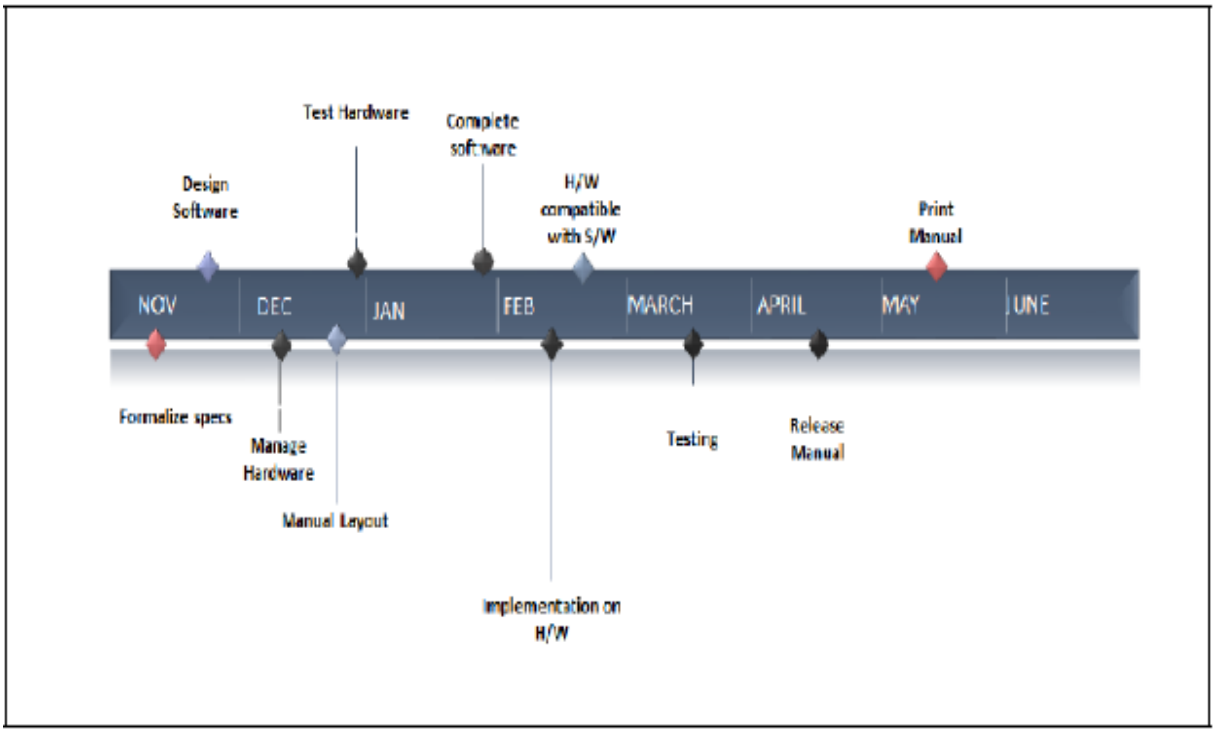


Figure 5.2: Timeline Chart

5.4 TEAM ORGANISATION

5.4.1 Team Structure

The team structure for the project is identified. There are total 4 members in our team and roles are defined. All members are contributing in all the phases of project.

5.4.2 Management Reporting and Communication

Mechanisms for progress reporting and inter/intra team communication are identified as per assessment sheet and lab time table.

Chapter 6

Software Requirement Specification

6.1 INTRODUCTION

6.1.1 Purpose and Scope of Document

The software requirement specification of our project will have the requirements which will be a baseline of our project. The software requirement specification will incorporate functional and non-functional requirements, system architecture, data flow diagrams, UML diagrams, experimental setup requirements and performance metrics.

A software requirements specification (SRS) is a document that is created when a detailed description of all aspects of the software to be built must be specified before the project is to commence. It is important to note that a formal SRS is not always written. There are many instances in which effort expended on an SRS might be better spent in other software engineering activities.

6.1.2 Overview of responsibilities of Developer

1. To have better understanding of the problem statement.
2. To know what are the hardware and software requirements of the proposed system.
3. To have understanding of the proposed system.
4. To plan various activities with the help of planner.
5. Designing, programming, testing etc.

6.2 USAGE SCENARIO

This section provides various usage scenarios for the system to be developed.

6.2.1 User Profiles

The profiles of all the user categories are described here.
User : The user will be viewer of the college website.
System : Handle the query. Provide the response regarding the input.

6.2.2 Use-cases

All use-cases for the software are presented. Description of the main Use cases using use case template is to be provided.

6.2.3 Use Case View

A use case diagram is a graphical representation of a user's interaction with the system and depicting the specifications of a use case. A use case diagram can show the different types of users of a system and the various ways in which they interact with the system.

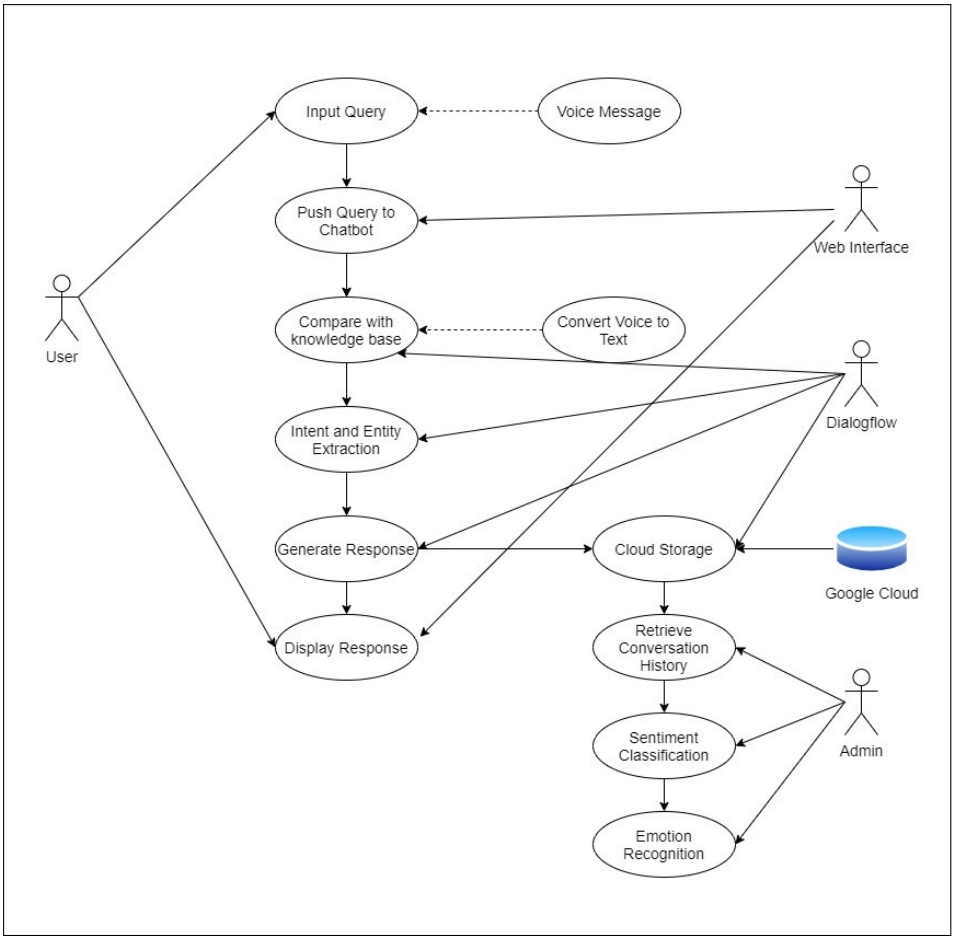


Figure 6.1: Use case diagram

6.3 DATA MODEL AND DESCRIPTION

6.3.1 Data Description

Input data will be the query typed in text based format or entered through voice based format.

6.4 FUNCTIONAL MODEL AND DESCRIPTION

Functional requirements define the internal workings of the software: that is, the technical details, data manipulation and processing and other specific functionality that show how the use cases are to be satisfied. They are supported by non-functional requirements, which impose constraints on the design or implementation.

The functional requirement of this service application can be broadly categorized into following parts –

- Entry of data
- Response data
- Classify emotion

- Entry of data:
  The User shall enter the respective query within required context.

- Response data:-
  The user shall get the valid response once processing is done.

- Classify emotion:-
  The emotion and corresponding sentiment should be computed.

6.4.1 Data Flow Diagram

A data flow diagram (DFD) is a graphical representation of the flow of data through an information system, modelling its process aspects. A DFD shows what kinds of information will be input to and output from the system, where the data will come from and go to, and where the data will be stored.

Level 0 Data Flow Diagram

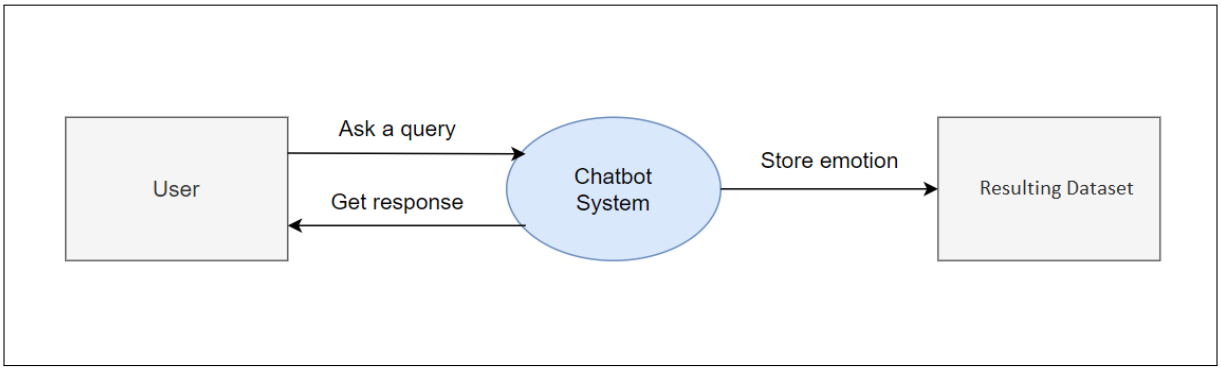


Fig 6.3 Level 0 Data Flow Diagram

Level 1 Data Flow Diagram

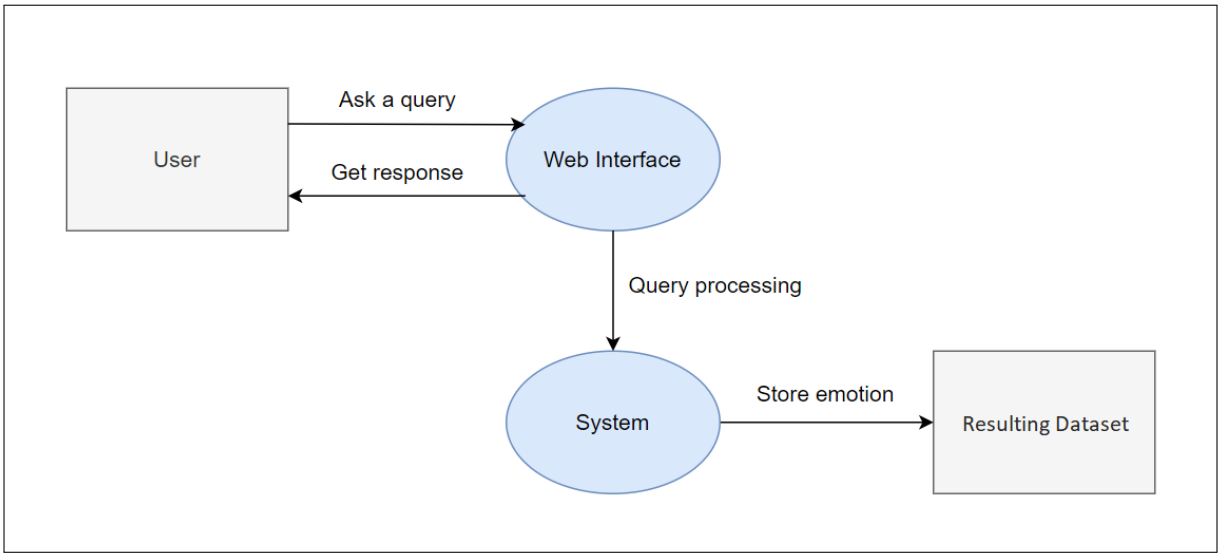


Fig 6.4 Level 1 Data Flow Diagram

6.4.2 Description of functions

- Access Application GUI : The user will be able to access the application GUI.
- The Response will be given : The user will be able to get the appropriate response of the query inserted.
- Display the result : The class of Object Detection will be displayed with the probability of the object.

6.4.3 Activity Diagram

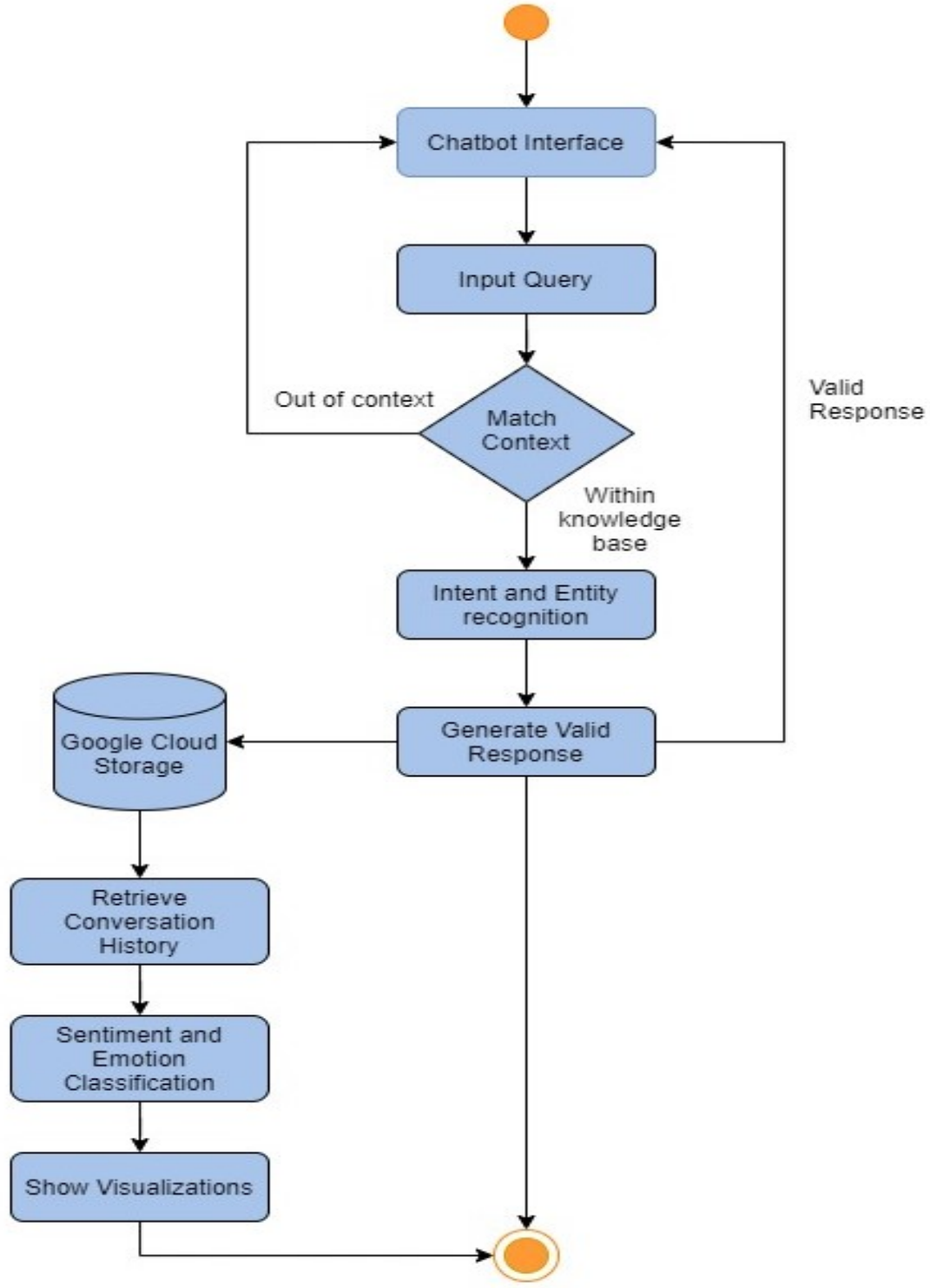


Figure 6.2: Activity Diagram

6.4.4 Non Functional Requirements:

1. Performance Requirements

- Reliability: If any exceptions occur during the execution of the software it should be caught and thereby prevent the system from crashing.
- Scalability: The system should be developed in such a way that new modules and functionalities can be added, thereby facilitating system evolution.
- Cost: The cost should be low because a free availability of software package.

2. Security Requirements :

All data will be encrypted using strong encryption algorithm and according to location encryption is done.

6.4.5 State Diagram

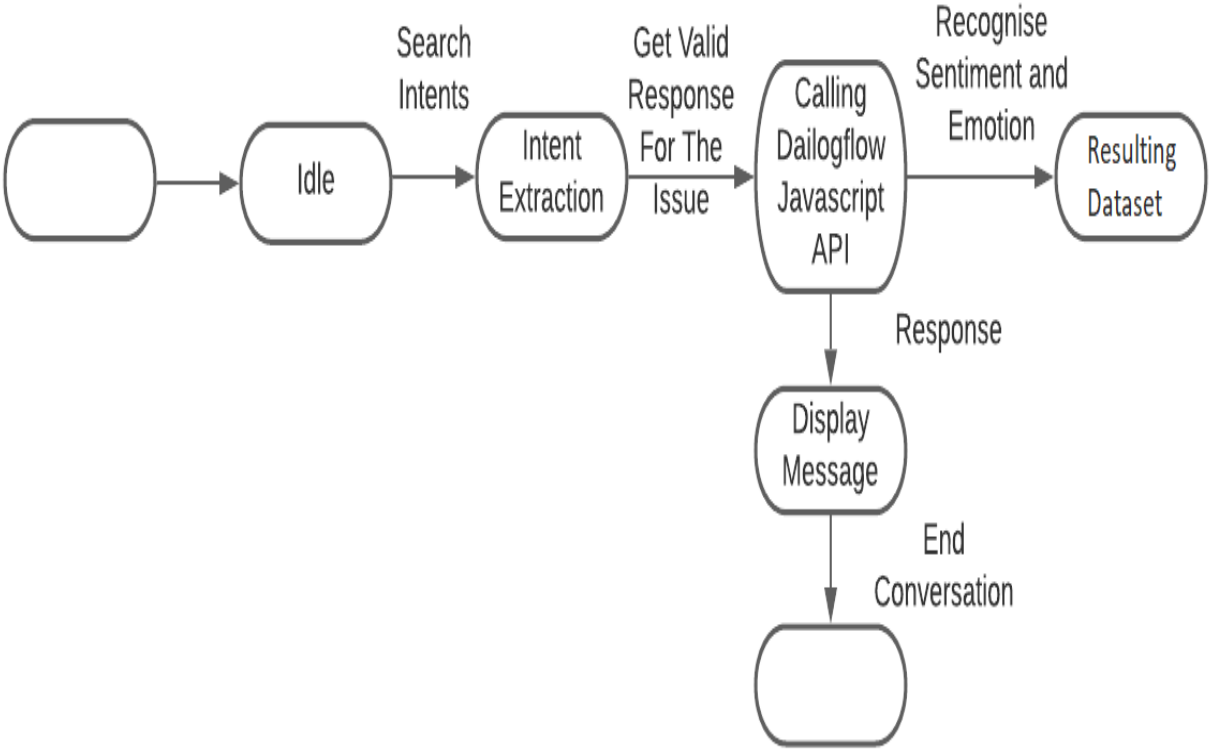


Figure 6.3: State Diagram

6.4.6 Design Constraints

The design constraints includes restrictions of the system that prevents smooth functioning of the application.

6.4.7 Software Interface Description

The software interface includes one component: browser window.

The browser window will display the result.

Chapter 7

Detailed Design Document using Appendix A and B

7.1 INTRODUCTION

This Application has four main basic modules, which gives feasibility to the Emotional Speech Recognition.

- System takes input from the user.
- Performs the processing of the query put forth by the user.
- Generate the response.
- Display the response to the user.

7.2 ARCHITECTURAL DESIGN

A description of the program architecture is presented. Subsystem design or Block diagram,Package Diagram,Deployment diagram with description is to be presented. The present invention has been summarized with the help of schematic artworks and overall internal details concerning inventing matter of subject.

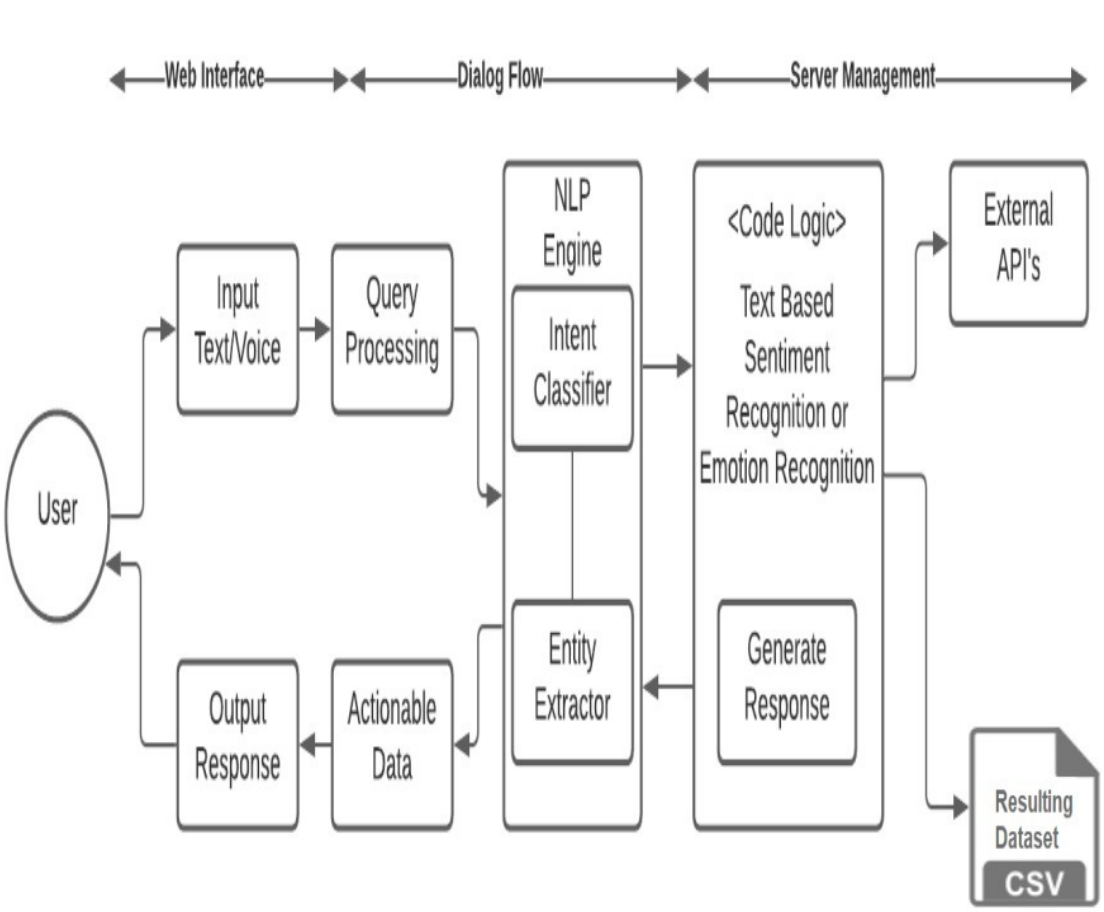


Figure 7.1: Architecture Diagram

7.3 DATA DESIGN

A description of all data structures including internal, global, and temporary data structures, database design (tables), file formats.

7.3.1 Internal software data structure

Majority of data passed among the components will be simple data type like textual based or voice based.

7.3.2 Global data structure

Data structured that are available to major portions of the architecture are described.

7.3.3 Temporary data structure

The data during training contributes to the temporary data structure.

7.4 COMPONENT DESIGN

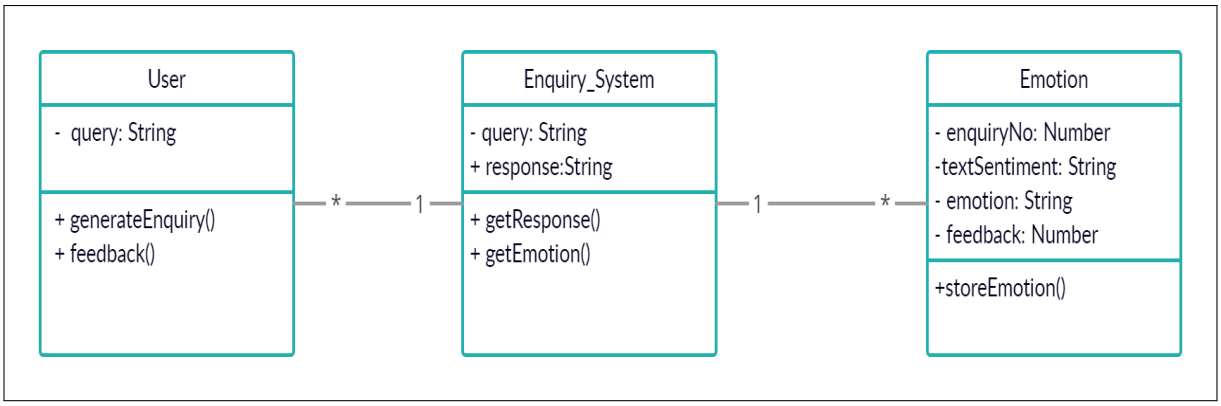


Figure 7.2: Class Diagram

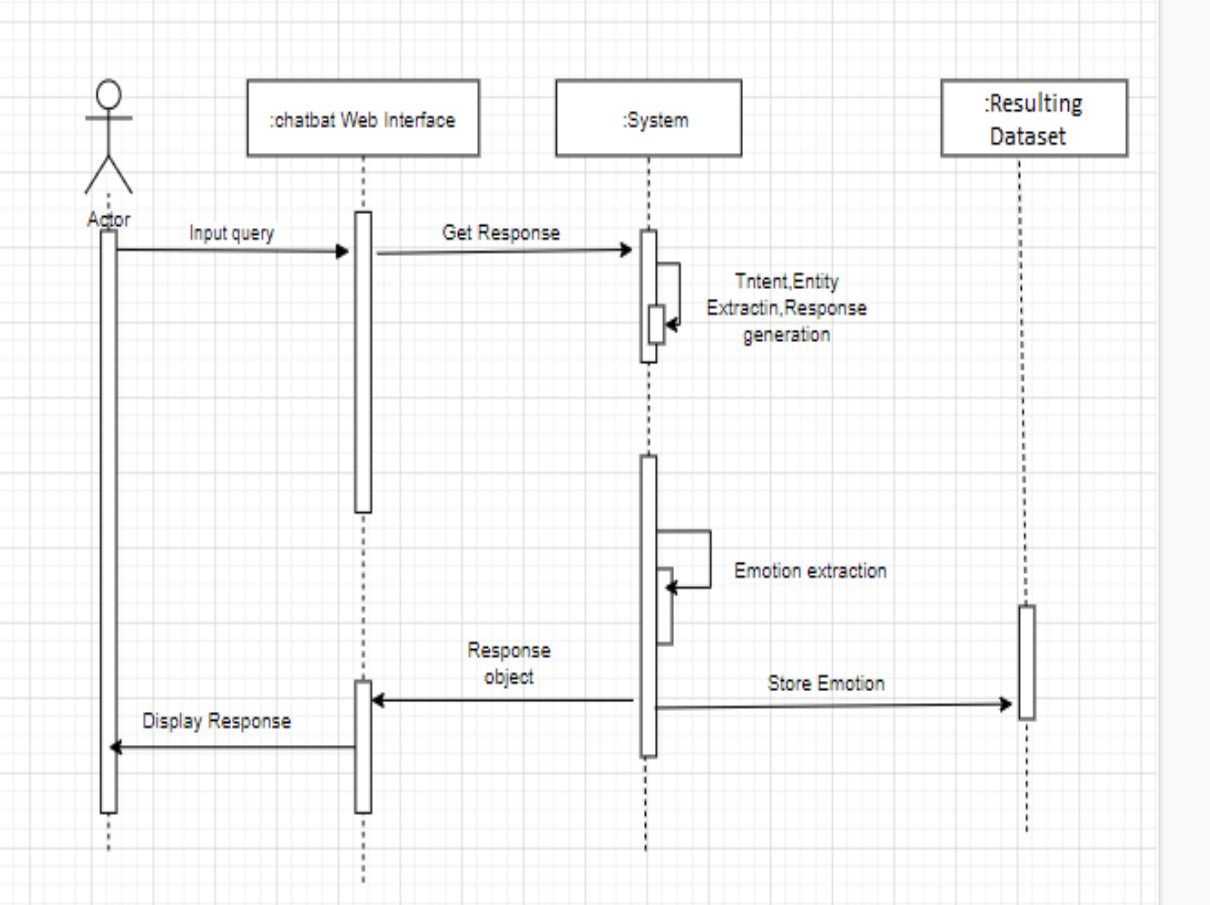


Figure 7.3: Sequence Diagram

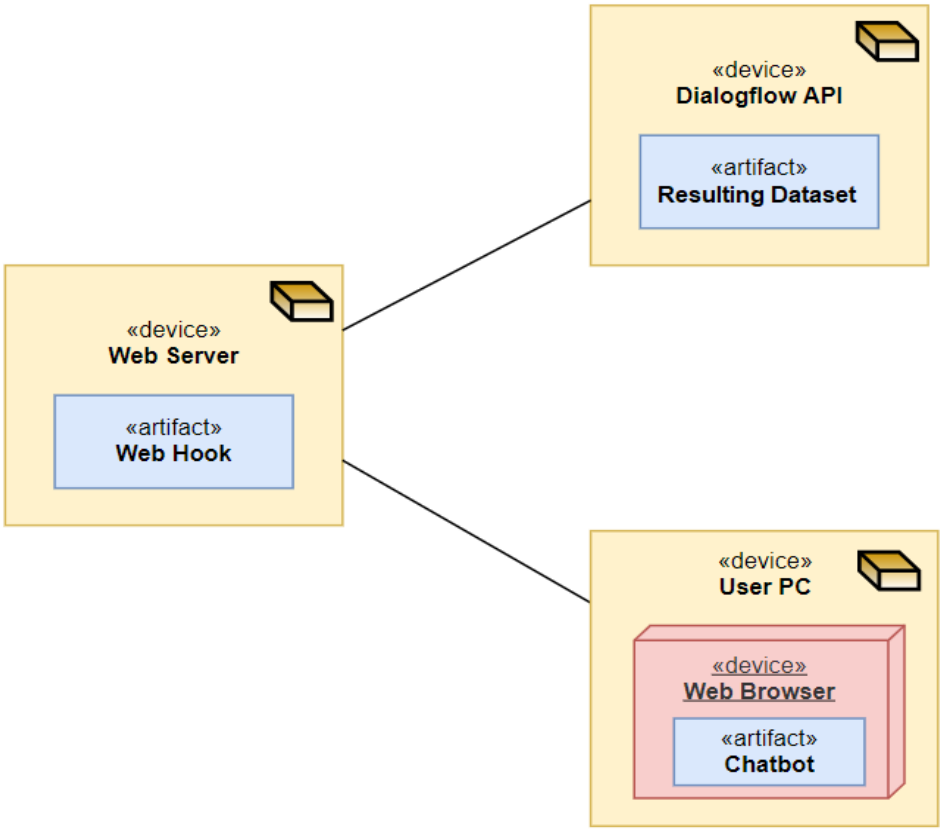


Figure 7.4: Deployment Diagram

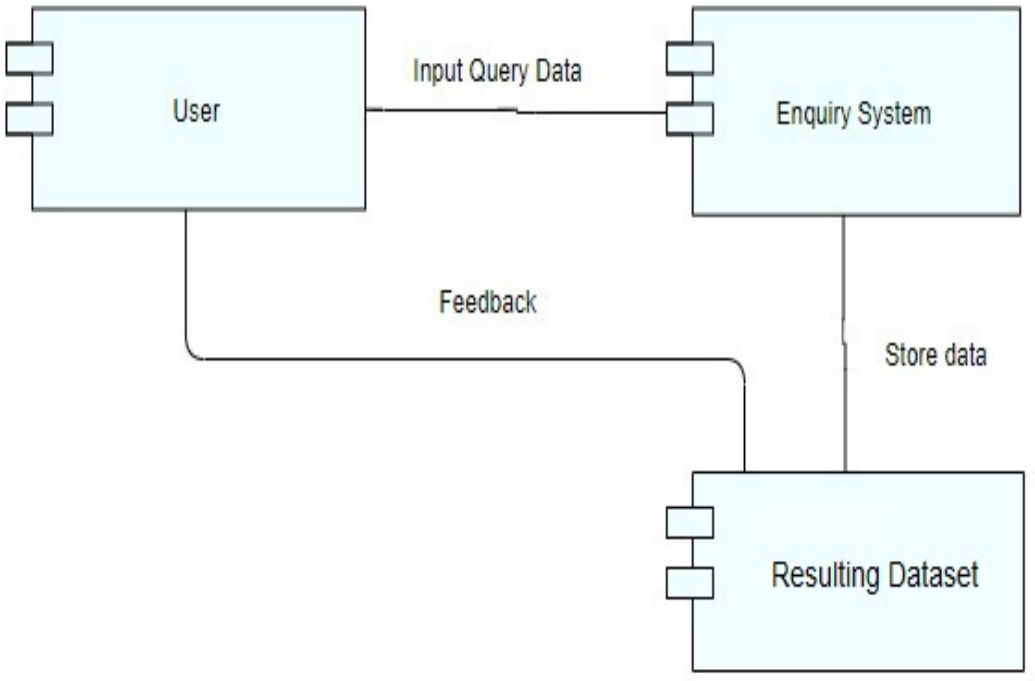


Figure 7.5: Component Diagram

Chapter 8

Project Implementation

8.1 INTRODUCTION

Chatbot experiences are becoming more advanced in these recent years. The ever extending advancements include emotion recognition, synthetic speech, language translations, etc. Enquiry through a bot within seconds is more convenient but not yet applied and practiced on a large scale by many people and organizations. College websites and other educational portfolios still retain greater response by the users and thus these are moving towards an automated process of interactive enquiry system.

Chatbot System :

In computer science, A chatbot is an artificial intelligence (AI) software that can simulate a conversation (or a chat) with a user in natural language through messaging applications, websites, mobile apps or through the telephone. through an algorithm. A chatbot is often described as one of the most advanced and promising expressions of interaction between humans and machines. However, from a technological point of view, a chatbot only represents the natural evolution of a Question Answering system leveraging Natural Language Processing (NLP). Formulating responses to questions in natural language is one of the most typical Examples of Natural Language Processing applied in various enterprises' end-use applications Image processing basically includes the following three steps:

- Accepting input query
- Query processing
- Output in which result is the response object of corresponding query.

There are three types of business chatbots namely, support, skills and assistant chatbots. Support chatbots are built to master a single domain, like knowledge about a company. Support chatbots need to have personality, multi-turn capability, and context awareness. Skills chatbots are typically more single-turn-type bots that do not require a lot of contextual awareness. Assistant chatbots are more or less a middle ground between the two bots above. They work best when they know a little bit about a variety of topics. Many people envision these bots will someday become navigators of all other bots that are out there now.

Figure 8.1: Emotionally Intelligent Chatbots

8.2 TOOLS AND TECHNOLOGIES USED

8.2.1 Python

Python is one of those rare languages which can claim to be both simple and powerful. You will find yourself pleasantly surprised to see how easy it is to concentrate on the solution to the problem rather than the syntax and structure of the language you are programming in.

Features of Python:-

- **Simple** :

  Python is a simple and minimalist language. Reading a good Python program feels almost like reading English. It allows you to concentrate on the solution to the problem rather than the language itself.

- **Easy to Learn** : As you will see, Python is extremely easy to get started with. Python has an extraordinarily simple syntax, as already mentioned.

- **Free and Open Source** : Python is an example of a FLOSS (Free/Libr´e and Open Source Software). In simple terms, you can freely distribute copies of this software, read its source code, make changes to it, and use pieces of it in new free programs. FLOSS is based on the concept of a community which shares knowledge. This is one of the reasons why Python is so good - it has been created and is constantly improved by a community who just want to see a better Python. High-level Language When you write programs in Python, you never need to bother about the low-level details such as managing the memory used by your program, etc.

- **Interpreted** : This requires a bit of explanation. A program written in a compiled language like C or C++ is converted from the source language i.e. C or C++ into a language that is spoken by your computer (binary code i.e. 0s and 1s) using a compiler with various flags and options. When you run the program, the linker/loader software copies the program from hard disk to memory and starts running it. Python, on the other hand, does not need compilation to binary. You just run the program directly from the source code. Internally, Python converts the source code into an intermediate form called byte codes and then translates this into the native language of your computer and then runs it. All this, actually, makes using Python much easier since

you don't have to worry about compiling the program, making sure that the proper libraries are linked and loaded, etc. This also makes your Python programs much more portable, since you can just copy your Python program onto another computer and it just works!

8.2.2 Librosa

Librosa is a python package for music and audio analysis. It provides the building blocks necessary to create music information retrieval systems. We can use it for sound analysis and extracting features of soundfile like chroma, mfcc, mel, etc for effective emotion recognition on voice based samples of the input.

8.2.3 CSS

CSS stands for Cascading Style Sheets. It describes how HTML elements are to be displayed on screen, paper, or in other media. It saves a lot of work. It can control the layout of multiple web pages all at once. CSS is of three types namely:

- External CSS
- Inline CSS
- Internal CSS

8.2.4 Google Log Files(GCP)

Log Files are files containing detailed logs on who and what is making requests to your website server. Every time a bot makes a request to your site, data (such as the time, date IP address, user agent, etc.) is stored in this log

8.3 METHODOLOGIES DETAILS

8.3.1 Query Processing

The input query entered by the user is processed as follows :

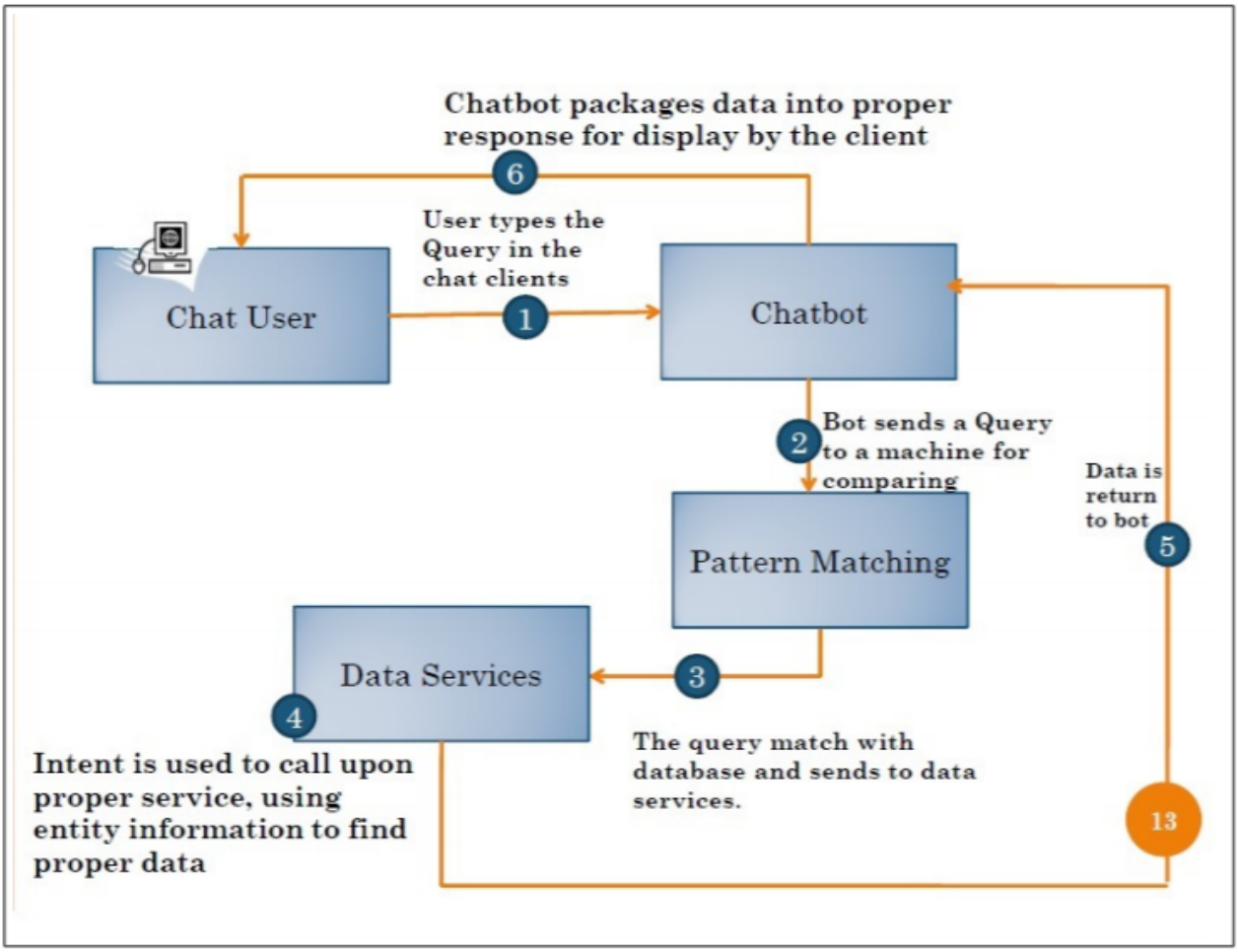


Figure 8.2: System Architecture Diagram

8.3.2 Pre-processing

It is a data mining technique that transforms raw data into an understandable format. Raw data (real world data) is always incomplete and that data cannot be sent through a system. That would cause certain errors. That is why we need to pre-process data before sending through a chatbot system. The input data can be typed incorrectly or can be semantically incorrect or in different variations. To tackle this, proper query processing has to be done.

Pre-processing techniques used in the proposed system

1. Intent Classification:- The intents of the query are matched with the intents in the database and context of the query is identified.
2. Entity Extraction:- Required entities are extracted for proper response generation.

8.3.3 Features

1. **EMOTION RECOGNITION :** Emotion recognition is the process of identifying human emotion. People vary widely in their accuracy at recognizing the emotions of others. Use of technology to help people with emotion recognition is a relatively nascent research area. Generally, the technology works best if it uses multiple modalities in context. Emotion recognition is used in society for a variety of reasons. Affectiva, which spun out of MIT, provides artificial intelligence software that makes it more efficient to do tasks previously done manually by people, mainly to gather facial expression and vocal expression information related to specific contexts where viewers have consented to share this information.

2. **SENTIMENT ANALYSIS :** Sentiment analysis (or opinion mining) is a natural language processing technique used to determine whether data is positive, negative or neutral. Sentiment analysis is often performed on textual data to help businesses monitor brand and product sentiment in customer feedback, and understand customer needs. Types of sentiment analysis are : Fine-grained Sentiment Analysis, Emotion detection, Aspect-based Sentiment Analysis, Multilingual sentiment analysis, etc.

3. **RESPONSE GENERATION :** Once the chatbot understands the user's message, the next step is to generate a response. One way is to generate a simple static response. Another way is to get a template based on intent and put in some variables. Developing successful chatbots is a non-trivial endeavor. In particular, the creation of high-quality natural language responses for chatbots remains a challenging and time-consuming task that often depends on high-quality training data and deep domain knowledge.

8.3.4 Methodology

The system is associated with a chat bot query process for each user. Every user will enter the input query and it will be passed to the query processing engine. This will enable the system to classify the context of input or meaning of the query for a machine. The same query will be passed for emotion recognition through text based or voice based mechanism. Valid response will be generated for the corresponding query and will be displayed on the web interface. If the query is spelled incorrect or semantically wrong then user will be prompted again to enter his query and response will not be generated.

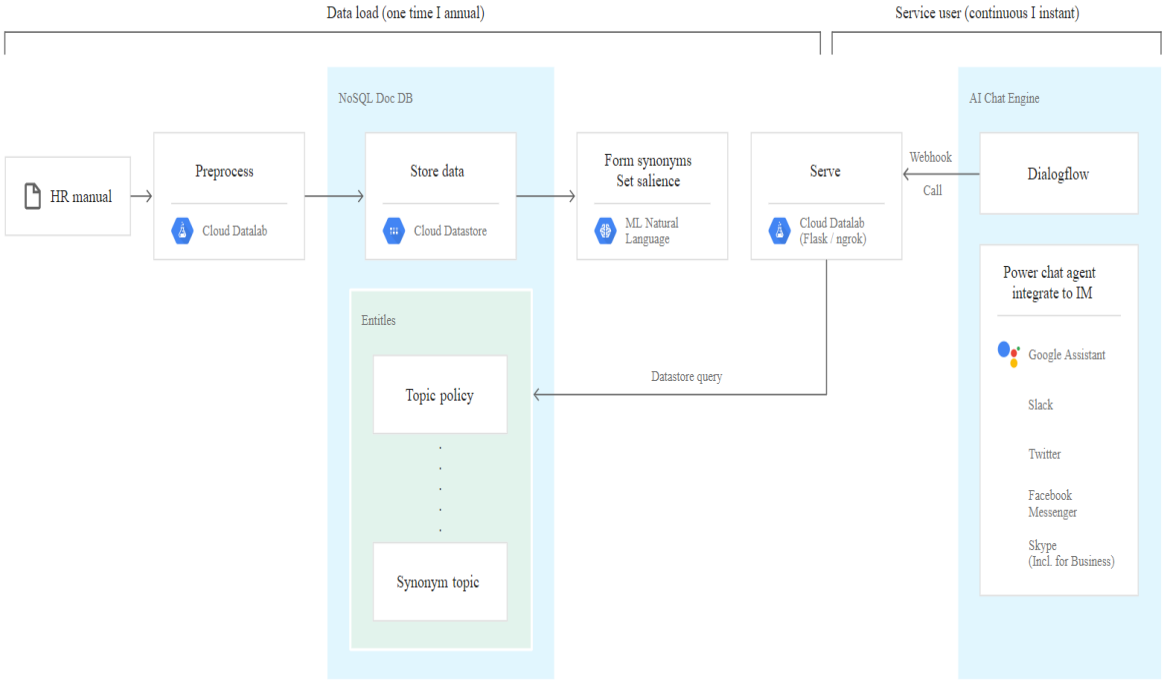


Figure 8.3: Dialogflow Working System

8.3.5 Output

Output obtained will be response for the query and it's emotion will be recorded accordingly.

The system is tested for various use cases in handling different forms of the same query. The system is feasible for trouble free interaction with the user in a natural human

to human conversation form as far as possible. The advanced technology with Machine Learning system permits the user to communicate without any restrictions or errors. Besides the convenience to the user, the system profits the website server by reducing the navigation overheads.

Algorithms

1. **Multinomial Naive Bayes Algorithm**
   (a) Multinomial Naive Bayes algorithm is a probabilistic learning method that is mostly used in Natural Language Processing (NLP)
   (b) The algorithm is based on the Bayes theorem and predicts the tag of a text such as a piece of email or newspaper article.
   (c) It calculates the probability of each tag for a given sample and then gives the tag with the highest probability as output
2. **Logistic Regression**
   (a) Logistic regression models the probability of the default class.
   (b) Logistic Regression is used when the dependent variable(target) is categorical.
   (c) For eg, To predict whether an email is spam (1) or (0), Whether the tumor is malignant (1) or not (0) etc.
3. **K-NN**
   (a) Classification happens by locating the object in the feature space, and comparing it with the k nearest neighbors (training examples). The majority vote decides on the classification.
   (b) k-nearest neighbors algorithm (k-NN) is a non-parametric machine learning method used for classification and regression.
   (c) k-NN is a type of instance-based learning, or lazy learning
4. **Sentiment Analysis**
   (a) Sentiment analysis is a predominantly classification algorithm aimed at finding an opinionated point of view and its disposition and highlighting the information of particular interest in the process.
   (b) It decides the sentiment of the input query based on the frequency of the positive, negative and neutral words in the input data.

8.4 VERIFICATION AND VALIDATION FOR ACCEPTANCE

8.4.1 Design

Design Validation

The first step in design validation is the design input given to the system. So here the input would be the input query of the user

The next step in design validation is design output. Here in this case the design output is expected to be the generated query response

Design Verification

Here the foremost step is developing a pre-trained system to validate the query correctness. Here the test cases are developed such that they act as input it gets compared to the expected intent type and form.

If any invalid results occur or if any procedures required modification, it is important to document the changes and get proper approval. Any issues are identified and logged as a defect at this stage.

The final step is creation of design verification report which gives the detailed summary of verification results https://www.overleaf.com/project/600858911d9a8176dc036197
which includes the configuration management, test results for each type of testing and issues found during the verification activity.

Chapter 9

Software Testing

9.1 FORMAL TECHNICAL REVIEWS

Technical reviews were conducted 4 times in each semester.The project was presented by all team members to the internal guide as well as an external guide. In the first technical review conducted, we had explained the objective,motivation and scope of the project. The algorithms were studied and compared and decided the possible output of the project that can be further improved. In this formal technical review, the panel members guided us by giving suggestions regarding the system. The reviews were:

1. Find accuracy for each model by providing test more classifiable intents.
2. Decide overall flow of system.

In the next technical review, the system module of the proposed system was explained. More proper explanation of the working and algorithm was explained along with the implementation. The following points were suggested

1. More number of test cases are needed to create.
2. Create more user friendly GUI.

The final review was finished with complete implementation and detailed explanation of the project and its applications.

SR.NO	Objective	Description	Observed Output	Expected Output	Result
1	To open chatbot on clicking the icon	In order to open the chatbot, click on the icon	Bot should be opened after clicking it.	The bot gets opened after clicking the icon	Pass
2	To display welcome message after howering into the chatbot	The message "Welcome To PCCOER" will be displayed.	Welcome message should be displayed after howering into the icon	The welcome message gets displayed after howering into the icon	Pass
3	To display default fallback intent if query is not processed	The message "Can you Rephrase it" will be displayed if query doesn't match.	The default fallback intent should be displayed if query doesn't match with the knowledge base	The default fallback message gets displayed if the query is not defined in the knowledge base	Pass

Figure 9.1: Test Cases table

Chapter 10

Results

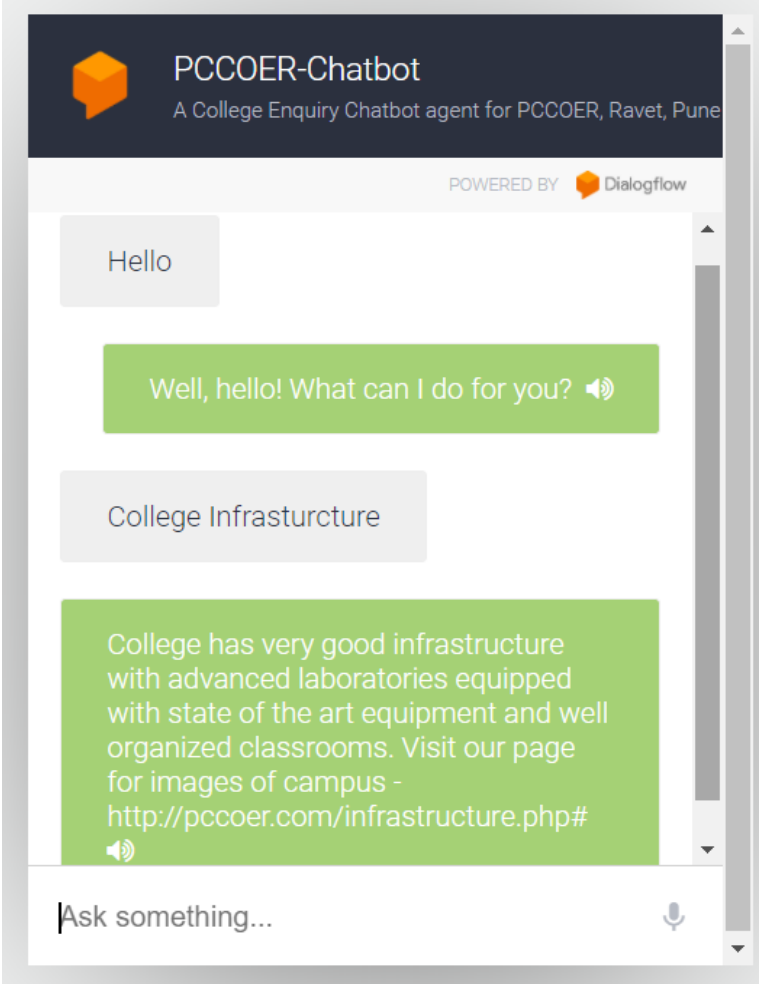


Figure 10.1: Bot Interface

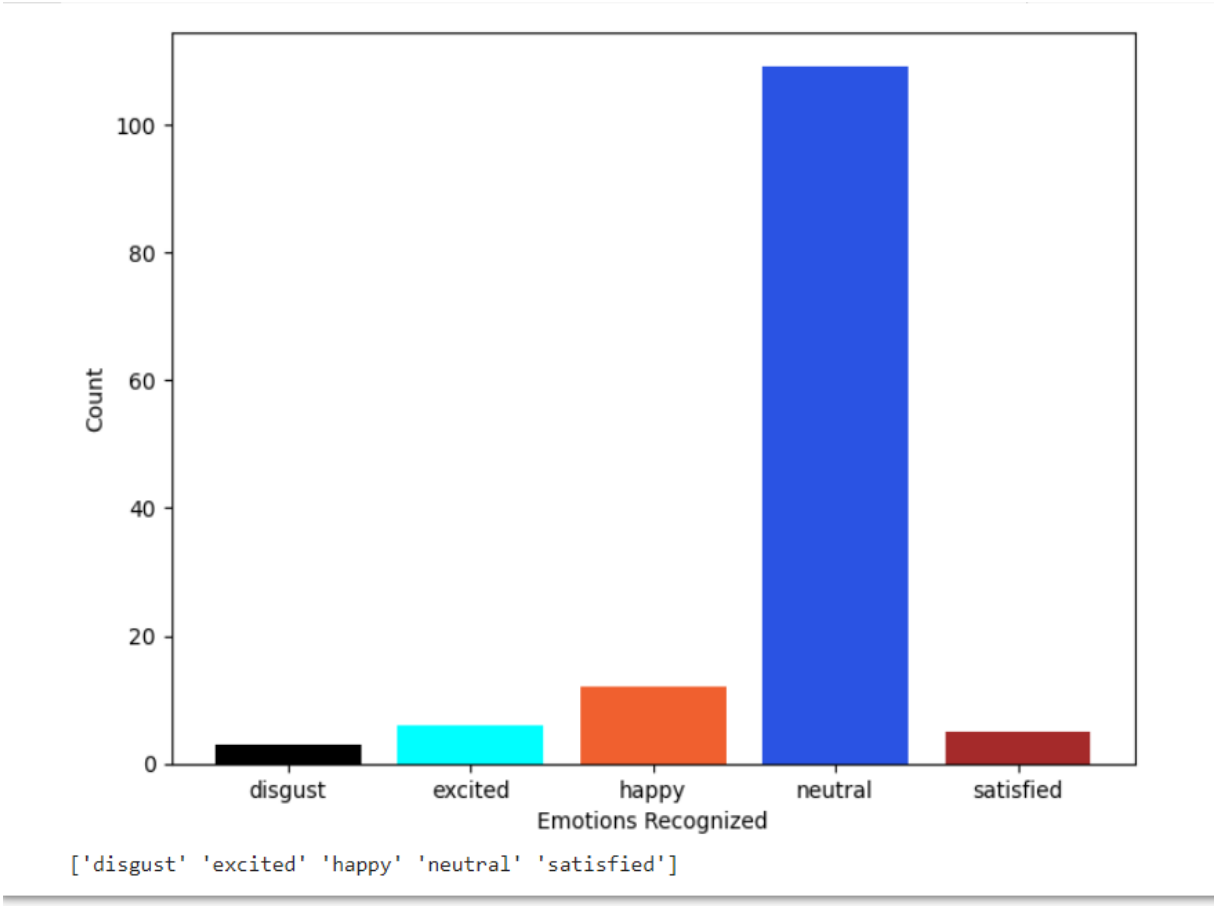


Figure 10.2: Polarity Wise Analysis Of Segments

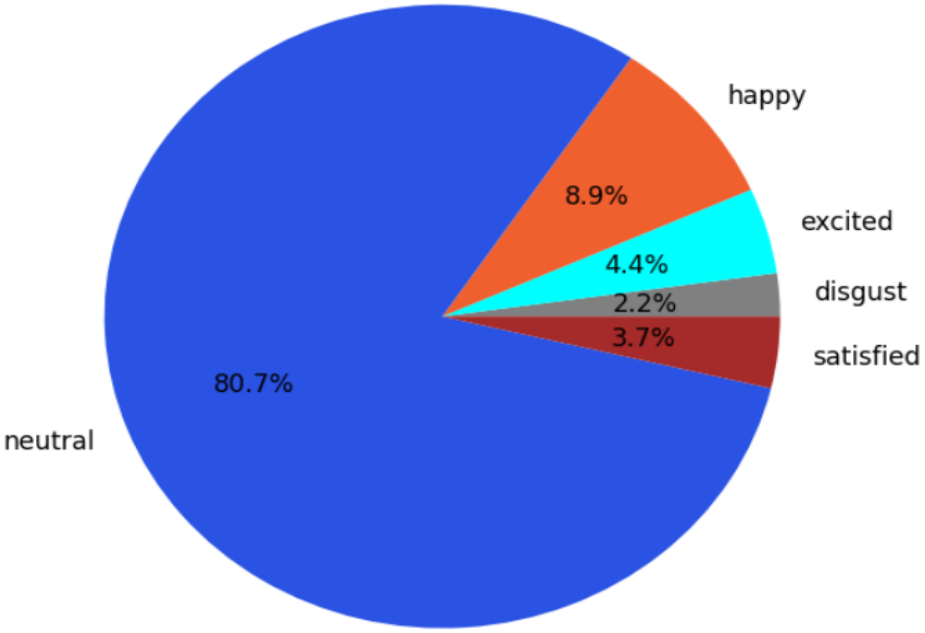


Figure 10.3: Pie Representation Of Segments Recognized

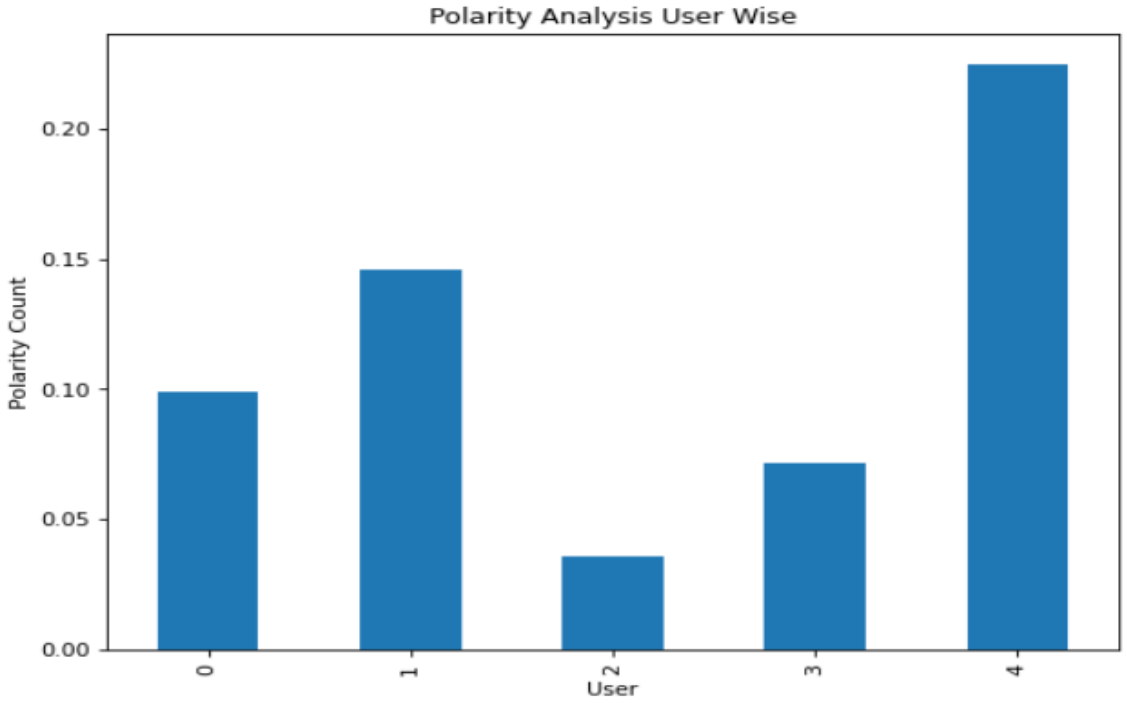


Figure 10.4: Polarity Wise User Segments

		timestamp	query	conversationId	Subjectivity	Polarity	Sentiment	emotion_recognized
2	0	2021-03-23T18:44:3	cancel	ABwppHFytFmvlLim(	0	0	Neutral	neutral
3	1	2021-03-23T18:44:3	cancel	ABwppHFytFmvlLim(	0	0	Neutral	neutral
4	2	2021-03-23T18:44:2	Do i need to take ori	ABwppHFytFmvlLim(	0.625	0.1875	Positive	neutral
5	3	2021-03-23T18:44:2	take original docume	ABwppHFytFmvlLim(	0.625	0.1875	Positive	neutral
6	4	2021-03-23T18:44:2	need to take original	ABwppHFytFmvlLim(	0.625	0.1875	Positive	neutral
7	5	2021-03-23T18:44:2	take original docume	ABwppHFytFmvlLim(	0.625	0.1875	Positive	neutral
8	6	2021-03-23T18:44:2	need to take original	ABwppHFytFmvlLim(	0.625	0.1875	Positive	neutral
9	7	2021-03-23T18:44:2	Do i need to take ori	ABwppHFytFmvlLim(	0.625	0.1875	Positive	neutral
10	8	2021-03-23T18:44:2	Do i need to take ori	ABwppHFytFmvlLim(	0.625	0.1875	Positive	neutral
11	9	2021-03-23T18:44:2	Is there any facility t	ABwppHFytFmvlLim(	0	0	Neutral	excited
12	10	2021-03-23T18:44:2	Is there any facility t	ABwppHFytFmvlLim(	0	0	Neutral	excited
13	11	2021-03-23T18:44:2	Is there any facility t	ABwppHFytFmvlLim(	0	0	Neutral	excited
14	12	2021-03-23T18:44:1	Can i apply online fc	ABwppHFytFmvlLim(	0	0	Neutral	neutral
15	13	2021-03-23T18:44:1	Can i apply online fc	ABwppHFytFmvlLim(	0	0	Neutral	neutral
16	14	2021-03-23T18:44:1	Can i apply online fc	ABwppHFytFmvlLim(	0	0	Neutral	neutral
17	15	2021-03-23T18:44:0	What documents are	ABwppHFytFmvlLim(	0	0	Neutral	excited
18	16	2021-03-23T18:44:0	what are the docume	ABwppHFytFmvlLim(	0	0	Neutral	neutral
19	17	2021-03-23T18:44:0	documents required	ABwppHFytFmvlLim(	0	0	Neutral	neutral
20	18	2021-03-23T18:44:0	documents required	ABwppHFytFmvlLim(	0	0	Neutral	neutral
21	19	2021-03-23T18:44:0	what documents req	ABwppHFytFmvlLim(	0	0	Neutral	neutral

Figure 10.5: Final Emotion Recognized In Resulting Dataset

Chapter 11

Deployment and Maintenance

11.1 INSTALLATION

11.1.1 Processing

1. Need Dialogflow in the chrome browser.
2. Python Libraries for core logic.
3. Webhook for calling web service api.
4. We'll also be creating a User-interface where the customers can access it on web.
5. We will store emotion in the form of CSV file.

Chapter 12

Conclusion

In this project, we proposed a chatbot system where the user can ask their college related queries in textual or voice based format . The query is then passed through an NLU engine through intent classifier and entity extractor and a valid response is generated and displayed to the user. The proposed methodology classifies emotions of users based on their input queries and also computes the overall sentiment. In the proposed work, data preprocessing is carried out using techniques such as count vectorization and lemmatization.
The overall accuracy achieved by the system is 74%.

Chapter 13

References

1. Ms.Ch.Lavanya Susanna, R.Pratyusha, P.Swathi, P.Rishi Krishna, V.Sai Pradeep," COLLEGE ENQUIRY CHATBOT", International Research Journal of Engineering and Technology (IRJET),on 2020

2. A. Harimi1*, A. Shahzadi1, A.R. Ahmadyfard2 and Kh.Yaghmaie Departmentof Electrical Engineering and Robotics, Shahrood University of technology, Iran09 February 2013

3. Sumit Wailthare, Tushar Gaikwad, Ketkee Khadse, Pooja Dubey,"Artificial Intelligence Based Chat-Bot",International Research Journal of Engineering and Technology (IRJET),on 2018

4. Gustavo Assuncao, Paulo Menezes, Fernando Perdigao, "Speaker Awareness for Speech Emotion Recognition", International Journal Of Online And Biomedical Engineering, 2020

Annexure A

Reviewers Comments of Paper Submitted

A.1 RESEARCH PAPER

1. Paper Title: Emotionally Intelligent Chatbot
2. Name of the Conference/Journal where paper submitted : Journal Of University Of Shanghai For Science And Technology (JUUST)
3. Paper accepted/rejected : Accepted And Published
4. Review comments by reviewer : None
5. Corrective actions if any : None

Annexure B

Laboratory Assignments on Project Quality Reliability Testing of Project Design

Study of required libraires needed for execution.

B.1 ASSIGNMENT 1

B.1.1 Guide Reviewes

1. Design the agent of proposed system.
2. Study the algorithm and software compatibility.
3. Decide the flow of application.

B.1.2 Actions Taken

Accepting User Query And generating relevent response

Based on the query, sentiments will be generated

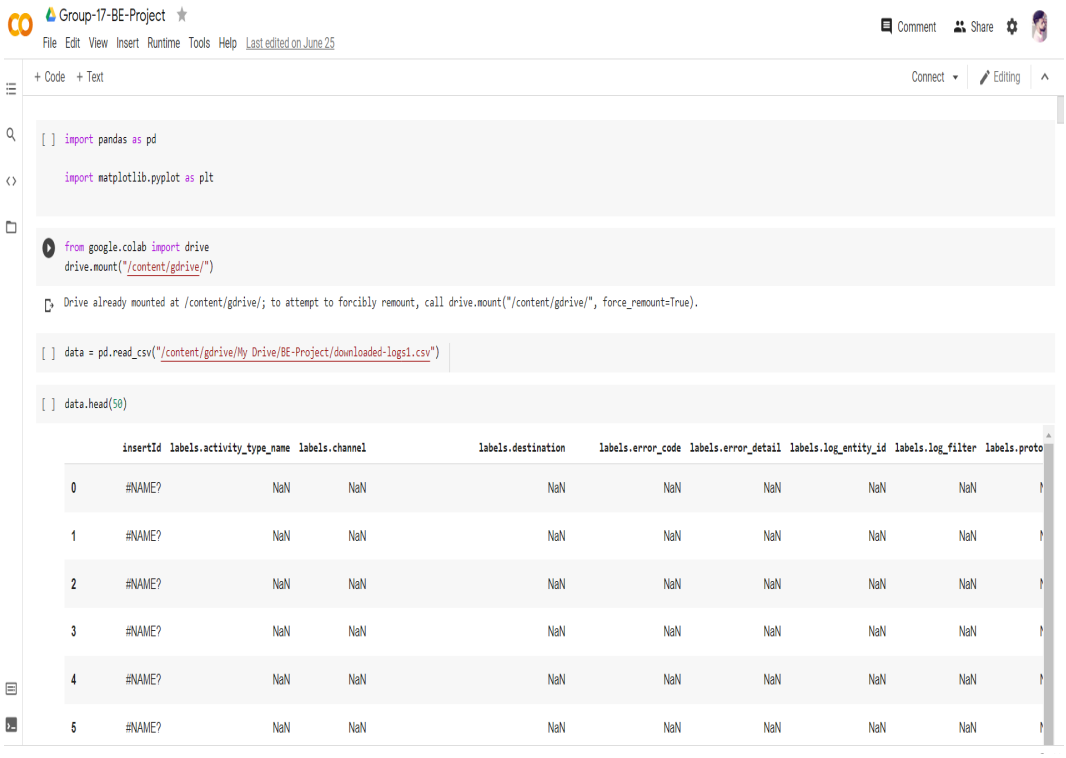
Finally, generating the recognised emotions and storing it in resulting dataset

B.2 ASSIGNMENT 2

Programming of project functions,training the modules, interfaces and GUI.

The entire project consists of three main tasks:

- Chatbot Interaction
- Sentiment Analysis
- Emotion Classification



	insertId	labels.activity_type_name	labels.channel	labels.destination	labels.error_code	labels.error_detail	labels.log_entity_id	labels.log_filter	labels.proto
0	#NAME?	NaN	NaN	NaN	NaN	NaN	NaN	NaN	N
1	#NAME?	NaN	NaN	NaN	NaN	NaN	NaN	NaN	N
2	#NAME?	NaN	NaN	NaN	NaN	NaN	NaN	NaN	N
3	#NAME?	NaN	NaN	NaN	NaN	NaN	NaN	NaN	N
4	#NAME?	NaN	NaN	NaN	NaN	NaN	NaN	NaN	N
5	#NAME?	NaN	NaN	NaN	NaN	NaN	NaN	NaN	N

Figure B.1: Code Screenshot

B.3 ASSIGMENT 3

Testing of various test cases for the project performed and generate various testing results.

SR.NO	Objective	Description	Observed Output	Expected Output	Result
1	To open chatbot on clicking the icon	In order to open the chatbot, click on the icon	Bot should be opened after clicking it.	The bot gets opened after clicking the icon	Pass
2	To display welcome message after howering into the chatbot	The message “Welcome To PCCOER” will be displayed.	Welcome message should be displayed after howering into the icon	The welcome message gets displayed after howering into the icon	Pass
3	To display default fallback intent if query is not processed	The message ”Can you Rephrase it” will be displayed if query doesn’t match.	The default fallback intent should be displayed if query doesn’t match with the knowledge base	The default fallback message gets displayed if the query is not defined in the knowledge base	Pass

Figure B.2: Test Cases table

Annexure C

Laboratory Assignment on Project Analysis of Algorithmic Design

C.1 ALGORITHM

C.1.1 Multinomial Naive Bayes Algorithm

(a) Multinomial Naive Bayes algorithm is a probabilistic learning method that is mostly used in Natural Language Processing (NLP)

(b) The algorithm is based on the Bayes theorem and predicts the tag of a text such as a piece of email or newspaper article.

(c) It calculates the probability of each tag for a given sample and then gives the tag with the highest probability as output

Multinomial Naive Bayes Emotion Classification Algorithm steps–

1. Take the input from the count vectorizer.
2. Take the input from the target column ‘emotion’.
3. Train the model.
4. Comparison of the features from test data to training data is done in following steps for classifying the emotion in the test data:

(i) Naive Bayes creates a frequency table of training data and lists count of all the words against corresponding emotion.

(ii) It finds probability for each emotion condition and creates a likelihood table.
(iii) It assumes that the effect of predictor (A) on a given class (B) is independent of the values of other predictors.
(iv) It computes the posterior probability for each emotion using the Naive Bayes theorem.
(v) The emotion with the highest probability will be the outcome of that word in the data.
(vi) Classified emotion of the string is returned.
5. Repeat steps from 1 to 4 until all the training set is computed.

C.1.2 Logistic Regression

(a) Logistic regression models the probability of the default class.

(b) Logistic Regression is used when the dependent variable(target) is categorical.

(c) For eg, To predict whether an email is spam (1) or (0), Whether the tumor is malignant (1) or not (0) etc.

C.1.3 K-NN

(a) Classification happens by locating the object in the feature space, and comparing it with the k nearest neighbors (training examples). The majority vote decides on the classification.

(b) k-nearest neighbors algorithm (k-NN) is a non-parametric machine learning method used for classification and regression.

(c) k-NN is a type of instance-based learning, or lazy learning

Annexure D

Plagiarism Report

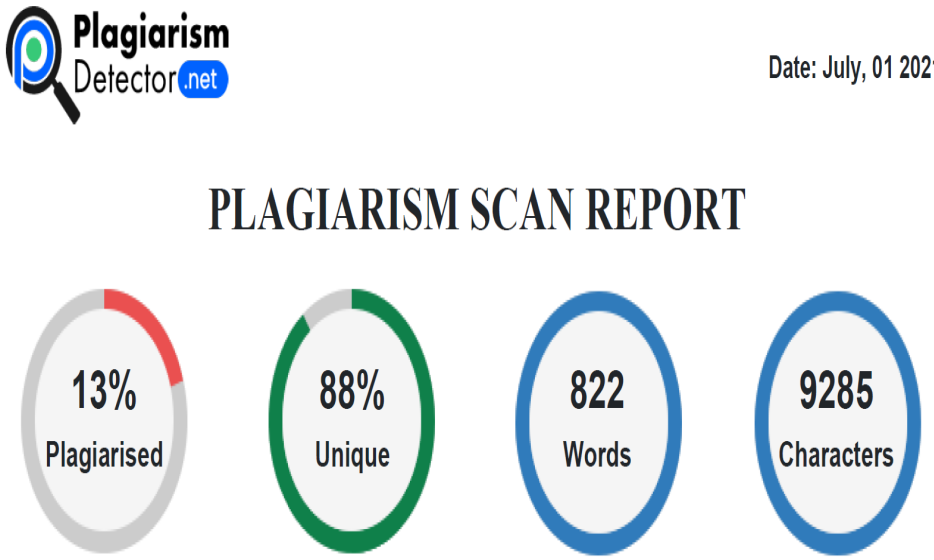


Figure D.1: Plagiarism Report

Annexure E

Project Planner

E.1 PROJECT TASK SET

Here in project planner for the project where in we have roughly estimated the time required for each phase and also the time required for entire project execution.
Major Tasks in the Project stages are:

- Task 1: Searching and selection of domain and topic thus building an abstract idea of a problem.
- Task 2: Project confirmation
- Task 3: Literature survey
- Task 4: Requirement analysis and gathering
- Task 5: Designing and model evaluation
- Task 6: Testing
- Task 7: Result Analysis

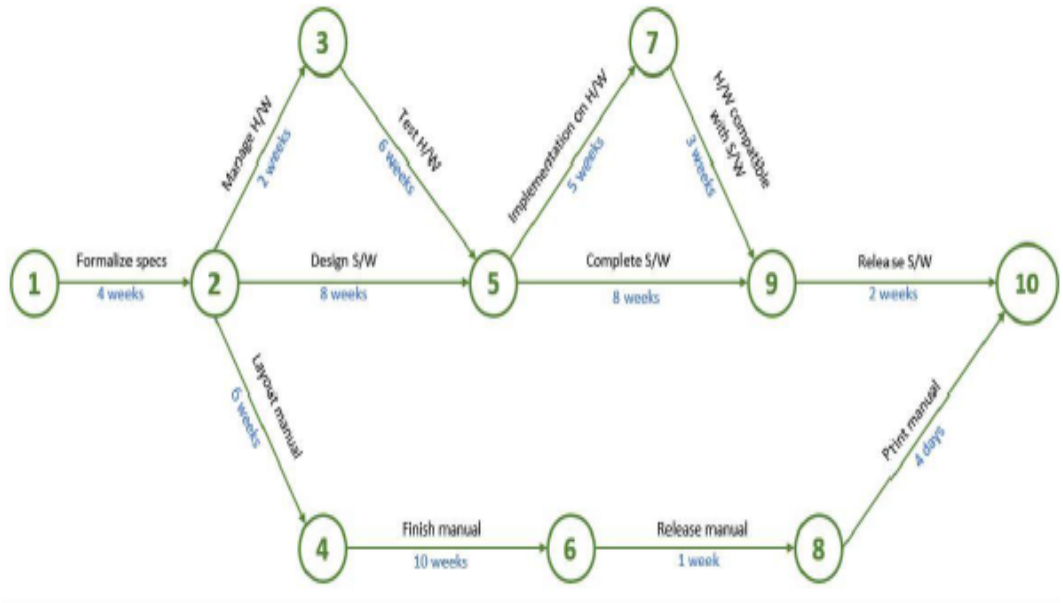


Figure E.1: Project Planner

Annexure F

Information of Project Group Members

1. Name: Shubham Deepak Kokane
   Date of Birth: 21/11/1998
   Gender: Male
   Permanent Address: 203/B2 Manish Garden, Udhyam Nagar, Pimpri Pune 411018
   Email: shubhamkkn7@gmail.com
   Contact No: +91 9823417110
   Placement Details: Birlasoft Pvt. Ltd.

2. Name:Shreeyash Ravindra Khalate
   Date of Birth: 24/05/1999
   Gender: Male
   Permanent Address: A/P- Pandare, Baramati, Pune 413110
   Email: shreekhalate555@gmail.com
   Contact no: +91 7066143707
   Placement Details: Pursing Higher Education.

3. Name: Sakshi Girish Dubewar

   Date of Birth: 20/01/2000

   Gender: Female

   Permanent Address: 10, Ratan-Rekha, Vasvi Nager, Pusad, 445204

   Email: sakshidubewar@gmail.com

   Contact no: +91 7741033104

   Placement Details: TCS

4. Name:Shreya Nilesh Newale

   Date of Birth: 28/06/1999

   Gender: Female

   Permanent Address: ”Sukhakarta”, Opp Mahalaxmi Temple, Laxmi Road, Chikhali, Pune-411062

   Email: shreyanewale11@gmail.com

   Contact no: +91 7028858686

   Placement Details: Pursing Masters