Mini Project Report

on

Voice Assistant Using Python

By

Bhupendra Huidrom (Reg. No.-201900422)

Aditya Deo Garg (Reg. No.-201900401)

Ayush Shekhar (Reg. No.-201900375)

Group Id – 24

In partial fulfilment of requirements for the award of degree in

Bachelor of Technology in Computer Science and Engineering

(2022)

Under the Project Guidance of

Dr. Nitai Paitya

Associate Professor, Dept of CSE

Sikkim Manipal Institute of Technology, Majitar

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

SIKKIM MANIPAL INSTITUTE OF TECHNOLOGY

(A constituent college of Sikkim Manipal University)

MAJITAR, RANGPO, EAST SIKKIM – 737136

PROJECT COMPLETION CERTIFICATE

This is to certify that the below mentioned students of Sikkim Manipal Institute of Technology have worked under my supervision and guidance from **17 January 2022 to 21 May 2022** and successfully completed the project entitled **"Voice Assistant Using Python"** in partial fulfilment of the requirements for the award of Bachelor of Technology in Computer Science and Engineering.

University Registration No	Name of Student	Course
201900422	**Bhupendra Huidrom**	**B. Tech (CSE)**
201900401	**Aditya Deo Garg**	**B. Tech (CSE)**
201900375	**Ayush Shekhar**	**B. Tech (CSE)**

Dr. Nitai Paitya

Associate Professor

Department of Computer Science and Technology

Sikkim Manipal Institute of Technology

Majitar, Sikkim – 737136

PROJECT REVIEW CERTIFICATE

This is to certify that the work recorded in this project report entitled **"Voice Assistant Using Python"** has been jointly carried out by **Bhupendra Huidrom (Reg. 201900422), Aditya Deo Garg (Reg. 201900401) and Ayush Shekhar (Reg. 201900375)** of Computer Science & Engineering Department of Sikkim Manipal Institute of Technology in partial fulfilment of the requirements for the award of Bachelor of Technology in Computer Science and Engineering. This report has been duly reviewed by the undersigned and recommended for final submission for Mini Project Viva Examination.

Dr. Nitai Paitya

Associate Professor

Department of Computer Science and Technology

Sikkim Manipal Institute of Technology

Majitar, Sikkim – 737136

CERTIFICATE OF ACCEPTANCE

This is to certify that the below mentioned students of Computer Science & Engineering Department of Sikkim Manipal Institute of Technology (SMIT) have worked under the supervision of **Dr. Nitai Paitya**, Associate Professor, Department of Computer Science and Engineering from **17 January 2022 to 21 May 2022** on the project entitled **"Voice Assistant Using Python"**.

The project is hereby accepted by the Department of Computer Science & Engineering, SMIT in partial fulfilment of the requirements for the award of Bachelor of Technology in Computer Science and Engineering.

University Registration No	Name of Student	Project Venue
201900422	**Bhupendra Huidrom**	**SMIT**
201900401	**Aditya Deo Garg**	**SMIT**
201900375	**Ayush Shekhar**	**SMIT**

Prof. (Dr.) Kalpana Sharma

Professor & Head of Department

Computer Science & Engineering Department

Sikkim Manipal Institute of Technology

Majitar, Sikkim – 737136

DECLARATION

We, the undersigned, hereby declare that the work recorded in this project report entitled "**Voice Assistant Using Python**" in partial fulfilment for the requirements of award of B. Tech (CSE) from Sikkim Manipal Institute of Technology (A constituent college of Sikkim Manipal University) is a faithful and bonafide project work carried out at "**SIKKIM MANIPAL INSTITUTE OF TECHNOLOGY**" under the supervision and guidance of **Dr. Nitai Paitya,** Associate Professor, Department of Computer Science and Engineering.

The results of this investigation reported in this project have so far not been reported for any other Degree or any other technical forum.

The assistance and help received during the course of the investigation have been duly acknowledged.

Bhupendra Huidrom (Reg. No.-201900422)

Aditya Deo Garg (Reg. No.-201900401)

Ayush Shekhar (Reg. No.-201900375)

ACKNOWLEDGEMENT

We take this opportunity to acknowledge indebtedness and a deep sense of gratitude to our guide **Dr. Nitai Paitya** whose valuable guidance and kind supervision gave us throughout the course which shaped the present work as it shows.

We pay our deep sense of gratitude to **Prof. (Dr.) Kalpana Sharma, H.O.D, Computer Science & Engineering Department, Sikkim Manipal Institute of Technology** for giving us the opportunity to work on this project and provide all support required.

We are obliged to our project coordinators **Dr. Sandeep Gurung and Mr. Biraj Upadhyaya** for elevating, inspiring and supervising the completion of our project.

We would also like to thank any other staff of the **Computer Science & Engineering Department, Sikkim Manipal Institute of Technology** for giving us continuous support and guidance that has helped us in completion of our project.

Bhupendra Huidrom (Reg. No.-201900422)

Aditya Deo Garg (Reg. No.-201900401)

Ayush Shekhar (Reg. No.-201900375)

DOCUMENT CONTROL SHEET

1	Report No	CSE/Mini Project/Internal/B.Tech/Section C/Group 24/2022
2	Title of the Report	Voice Assistant Using Python
3	Type of Report	Technical
4	Author	Bhupendra Huidrom, Aditya Deo Garg, Ayush Shekhar
5	Organizing Unit	Sikkim Manipal Institute of Technology
6	Language of the Document	English
7	Abstract	The voice assistance takes the voice input through our microphone (Bluetooth and wired microphone) and it converts our voice into computer understandable language, gives the required solutions and answers which are asked by the user.
8	Security Classification	General
9	Distribution Statement	General

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABSTRACT

In this contemporary generation, everyday existence became smarter and interlinked with technology. We already understand some voice assistants like Google, Siri and so forth. This project works on voice input and gives output via speech and shows the text at the display screen. The primary timetable of our voice assistance makes people clever and supply instant and computed results. The voice help takes the voice input thru our microphone (Bluetooth and wired microphone), and it converts our voice into computer understandable language, gives the specified answers and answers which can be asked by the consumer. This assistant connects with the world-wide internet to offer effects that the person has questioned. Natural language processing set of rules enables computer machines to have interaction in conversation using herbal human language in many forms.

In this project, we purpose to offer a one-prevent for speech popularity and a seek engine-type internet app the usage of voice instructions simplest. Then we purpose to add more features that use ML to create a single huge software.

CHAPTER 1: INTRODUCTION

1.1 General Overview

Nowadays the improvement of Artificial Intelligence (AI) structures which can organize a natural human-system interplay (thru voice, communication, gestures, facial expressions, and many others.) are gaining popularity. One of the maximum popular subjects is the direction of interaction, primarily based on the expertise of the system through the machine of the herbal human language. It's miles no longer a human who learns to communicate with a machine, but a machine learns to talk with a human, exploring his movements, habits, conduct and looking to be his personalised assistant.

Fig 1.1: Voice Assistant

Virtual assistants are software packages that assist you ease your everyday duties, which includes showing climate reports, growing remainders and many others. They can take instructions thru textual content (online chatbots) or by voice. Voice-based totally intelligent assistants need an invoking phrase to set off the listener, followed by way of the command. We have such a lot of digital assistants, which include apple's Siri, amazon's Alexa, and Microsoft's Cortana.

This machine is designed to be used efficaciously on computer systems. Personal assistance Software program improves user's productiveness by way of dealing with ordinary tasks of the consumer and by using

Supplying facts from a web supply to the user.

Internet searches performed thru cellular gadgets have overtaken the ones performed the usage of a PC and the analysts are already predicting that 50% of searches can be through voice through 2023.This project become started out on the premise that there's enough brazenly to be had facts and data on the net that can be utilized to build a digital assistant that has get right of entry to creating smart choices for recurring consumer sports.

1.2 Literature Survey

Sl No.	Author	Paper and Publication Details	Findings	Relevance to the project
1.	Rustam Shadiev,Wu-Yuin Hwang ,Nian-Shing Chen National Taiwan Normal University Yueh-Min Huang National Cheng Kung University	Review of Speech-to-Text Recognition Technology for Enhancing Learning Published - 2017 Aug 3 Publication Series- Proceedings - IEEE 17th International Conference on Advanced Learning Technologies, ICALT 2017	Creating an adequate accurate rate of strings from speech of speaker .	Improvement in Speech-to-text learning in the last few years.

2.	Asha Shajee, Dhruv Patel, Rahul Mishra, Himangshu Saikia, Narendran.M	A Survey: Speech Recognition Application Published in: International Journal of Advances in Electronics and Computer Science Volume 4, Issue 11 on Nov 2017	How to decipher the acoustic signals sent by the user which are recognized by the microphone and transform it into strings.	Speech-Recognition is the process by which the system understands user's spoken word and convert them to strings for the system to understand them.
3.	Suhas R. Mache , Manasi R. Baheti , C. Namrata Mahender	Review on Text-To-Speech Synthesizer. Published At International Journal of Advanced Research in Computer and Communication Engineering, Volume 4, Issue 8, August 2015	Text processing: It is a process where first a string is inspected, systematized and then converted into some linguistic representation. Speech Generation: Generation of	Using of TSS by computer to convert a text into audio using an artificial producer of human speech known as speech synthesizer which is used in both in software and hardware form. ·

			speech by using evaluation of phonetics	

Table 1.2.1: Literature Survey

1.3 Problem Definition

Generally, the consumer wishes to manually control multiple units of packages to finish one project. For instance, a consumer trying to make a tour plan wishes to check for airport codes for close by airports and then test travel web sites for tickets among combos of airports to attain the destination. There may be a want for a machine that could control obligations effortlessly. We already have a couple of digital assistants, but we hardly use it. There are several human beings who've troubles in voice recognition. These structures can apprehend English phrases; however, they fail to understand in our accessory. Our way of pronunciation is manner distinct from theirs.

•   There is a need of a virtual assistant that may recognize English in an Indian accessory and paintings on desktop device.
•   Whilst a digital assistant is not able to solution questions as it should be, it's because it lacks the right context or doesn't recognize the rationale of the question.

•   Its potential to answer questions relevantly simplest occurs with rigorous optimization, involving each people and device gaining knowledge of.

Continuously ensuring solid exceptional manage techniques will even assist manage the risk of the digital assistant mastering undesired horrific behaviours. They require large amount of information to be fed for it to paintings effectively. Virtual assistants must be capable of version complex task dependencies and use those models to propose optimized plans for the consumer. It wishes to be examined for finding surest paths when a mission has more than one sub duties and every sub duty will have its very own sub duties. In this type of case there may be multiple solutions to paths, and it ought to be capable of bear in mind consumer alternatives, other energetic responsibilities, priorities to endorse a selected plan

1.4 Solution Strategy

Speech popularity, or speech-to-textual content, is the ability of a device or program to pick out words spoken aloud and convert them into readable textual content. Rudimentary speech reputation software has a limited vocabulary and can best perceive words and terms whilst spoken virtually. More sophisticated software can take care of natural speech, one-of-a-kind accents, and numerous languages.

Speech recognition uses an extensive array of research in computer technology, linguistics, and computer engineering. Many current devices and text-targeted programs have speech reputation capabilities in them to allow for less complicated or hands-loose use of a device.

In python we use, speech recognition 3.8.1 library for speech popularity, which is the machine's ability to pay attention to the spoken words and convert them to textual content even as pyttsx3 module is used for the conversion of text to speech.
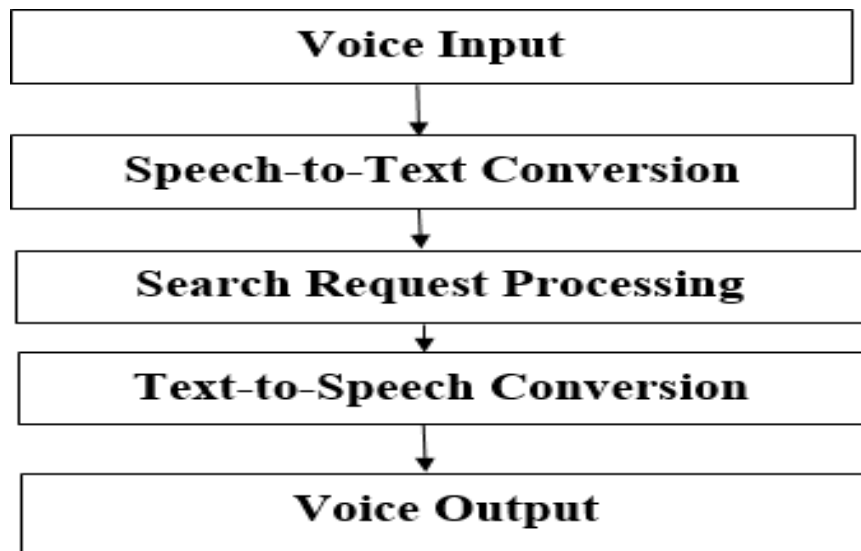
```
┌─────────────────────────────────────────┐
│              Voice Input                  │
└─────────────────────────────────────────┘
                    │
                    ▼
┌─────────────────────────────────────────┐
│        Speech-to-Text Conversion          │
└─────────────────────────────────────────┘
                    │
                    ▼
┌─────────────────────────────────────────┐
│        Search Request Processing          │
└─────────────────────────────────────────┘
                    │
                    ▼
┌─────────────────────────────────────────┐
│        Text-to-Speech Conversion          │
└─────────────────────────────────────────┘
                    │
                    ▼
┌─────────────────────────────────────────┐
│              Voice Output                 │
└─────────────────────────────────────────┘
```

Fig 1.4.1: Solution Strategy Part 1

A software application turns the sound a microphone fact into written language that computers and human beings can understand, following those four steps:

- examine the audio.

- damage it into elements

- digitize it into a pc-readable format; and

- use an algorithm to fit it to the most appropriate text representation inside the dataset.
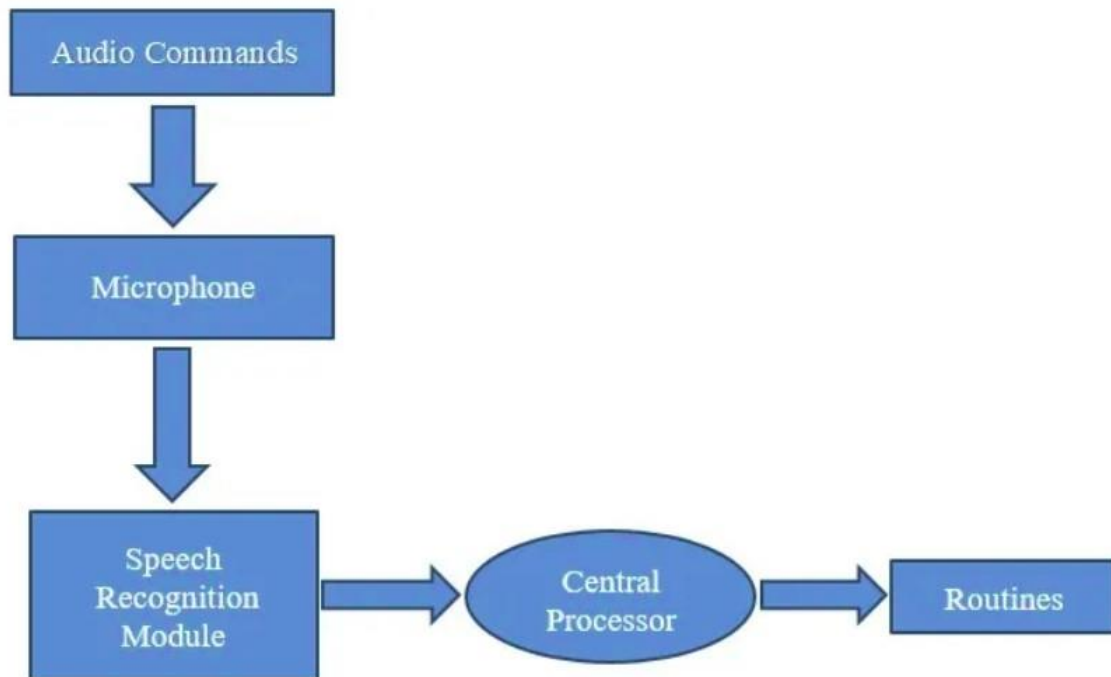
Fig 1.4.2: Solution Strategy Part 2

Software Requirements:

The software is designed to be mild weight so that it doesn't be a burden at the gadget strolling it. This gadget became build preserving in thoughts the usually to be had hardware and software program compatibility.

Hardware:

- Pentium-pro processor or later.
- RAM 512MB or more.

Software:
- Windows 7(32-bit) or above.
- Python 2.7 or later
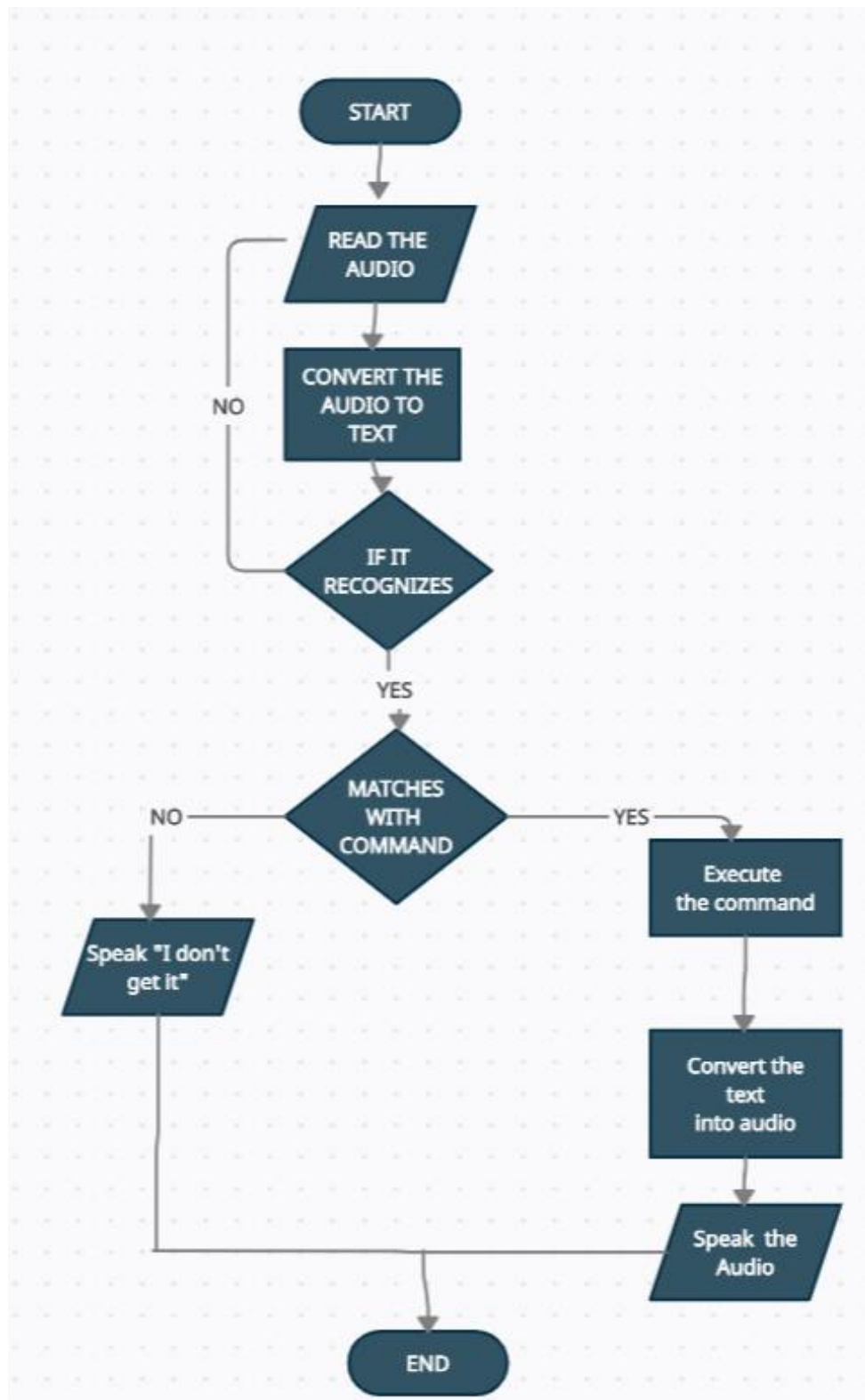- Chrome Driver

Python Libraries:
- Pyttsx3
- SpeechRecognition

1.5 Software Design

Fig 1.5.1: Flow Chart
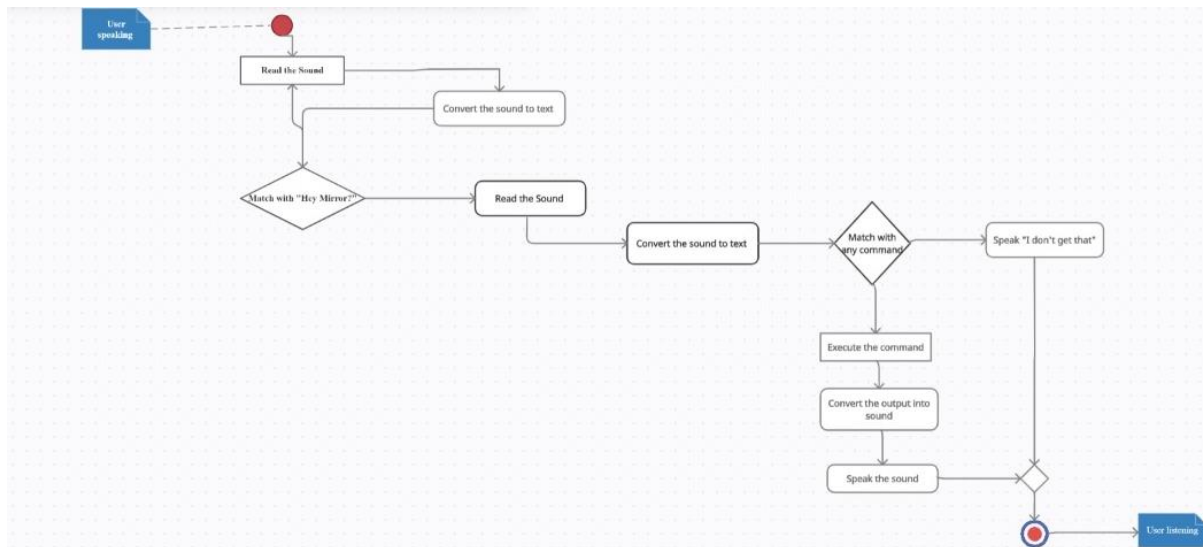
Fig 1.5.2: Activity Diagram

1.6 Organization of the Report

Chapter 1 Introduction gives the basic idea about the approach on how we solve the problems and the basic problem that we faced.

Chapter 2 Implementation talks about the details and technicalities of implementation.

Chapter 3 Limitation talks about the limitation that this project had and the future expansion that could be possible.

Chapter 4 Future Scope talks about the future scope of the project

Chapter 5 Gantt Chart discusses the Gantt Chart timeline of the project

Chapter 6 Conclusion discusses the conclusion of the project

Chapter 7 References discusses the references used for the project creation

CHAPTER 2: IMPLEMENTATION DETAILS

1. List of all the modules we need, to successfully run speech to text function and back text to speech function.

```
1   from urllib.parse import quote_plus
2   import pyttsx3 #pip install pyttsx3
3   import speech_recognition as sr #pip install speechRecognition
4   import datetime
5   import wikipedia #pip install wikipedia
6   import webbrowser
7   import os
8   import smtplib
9   import csv
10  import pywhatkit
11  from PIL import Image
12  # import pyjokescli
```

Fig 2.1: Imported Libraries

2. Made complete module for listening and recognizing the voice of user.

```python
def takeCommand():

    r = sr.Recognizer()
    #"""Recognizes the audio and sends it for display to displayText."""

    with sr.Microphone() as source:
        # use the default microphone as the audio source
        #listen for the first phrase and extract it into audio data
        print("Listening...")
        r.pause_threshold = 1
        r.adjust_for_ambient_noise(source)
        audio = r.listen(source)
        print (audio)

    try:
        print("Recognizing...")
        # received audio data, now need to recognize it
        query = r.recognize_google(audio, language='en-in')
        print(f"User said: {query}\n")

    except Exception as e:
        # print(e)
        print("Say that again please...")
        speak("Say that again please")
        return "None"
    return query
```

Fig 2.2: Modules for Recognising Speech

Fig 2.3: Output

3. Dataset containing different types of queries.

Fig 2.4: Dataset

CHAPTER 3: LIMITATIONS

•   Due to limited dataset this app cannot solve every question.

•   Mics with no diaphragm do no longer work effectively with this system because of defective background noise minimisation.

•   Protection problems like identity theft and many others can stand up.

•   Due to the nature of the tool, they should always listen to the user that may bring about violation of privacy.

CHAPTER 4: FUTURE SCOPE

Voice assistants will retain to provide greater individualized reviews as they get higher at differentiating among dissimilar sound though developers ought to cope with this complexity of understanding and analysing one-of-a-kind accents, exclusive slangs. They will also want to consciousness on retaining a person enjoy that is constant within the coming years as complexity will become extra of a challenge, as in upcoming years the visible interface will be available and the user can interact with them.
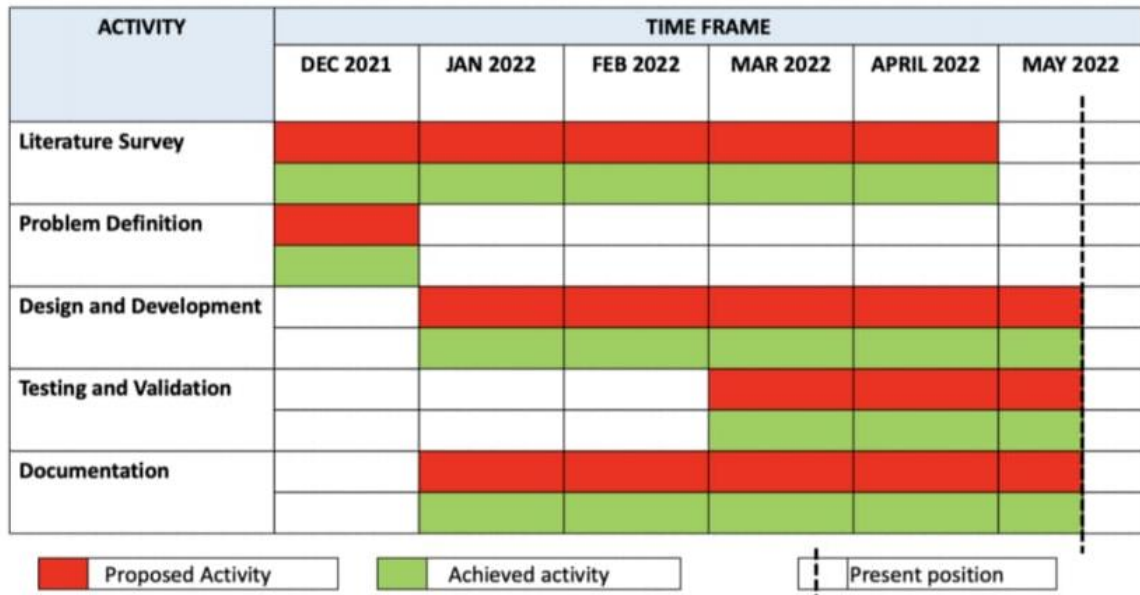
CHAPTER 5: GANTT CHART

ACTIVITY	TIME FRAME					
	DEC 2021	JAN 2022	FEB 2022	MAR 2022	APRIL 2022	MAY 2022
Literature Survey	🟥	🟥	🟥	🟥	🟥	
	🟩	🟩	🟩	🟩	🟩	
Problem Definition	🟥					
	🟩					
Design and Development		🟥	🟥	🟥	🟥	🟥
		🟩	🟩	🟩	🟩	🟩
Testing and Validation				🟥	🟥	🟥
				🟩	🟩	🟩
Documentation		🟥	🟥	🟥	🟥	🟥
		🟩	🟩	🟩	🟩	🟩

🟥 Proposed Activity   🟩 Achieved activity   ┊ Present position

Table 5.1: Gantt Chart

CHAPTER 6: CONCLUSION

On this project "Voice Assistant Using Python", we discussed the layout and implementation of digital assistance. The modular nature of this undertaking makes it greater bendy and less complicated to add extra features without disturbing modern system functionalities.

The main goal of the assignment was to provide a one stop solution for converting speech to text and supplying a voice primarily based digital assistant python app which could carry out many duties for the user. This voice based assistant help users in multitasking.

CHAPTER 7: REFERENCES

[1] Archana Balyan, S.S. Agrwal and Amita Dev, Speech Synthesis: Review, IJERT, ISSN 2278-0181 Vol. 2 (2013) p. 57 – 75

[2] Alistair Conkie, Thomas Okken, Yeon-Jun Kim, Giuseppe Di Fabbrizio, Building Text-To-Speech Voices in the Cloud, in Proc. AT&T Labs Research, Park Avenue, Florham Park, NJ-USA).

[3] Mark Tatham and Katherine Morton, Developments in Speech Synthesis (John Wiley & Sons, Ltd. ISBN: 0-470-85538-X, 2005)

[4] Mattingly I. G., Speech Synthesis for Phonetic and Phonological Models, T.A. Sebeok (Ed.) Current Trends in Linguistics, Vol. 12, (1974) p. 2451-2487