Voice Assistant using Speech Recognition

Submitted in partial fulfillment of the requirements of the degree

B.Tech.

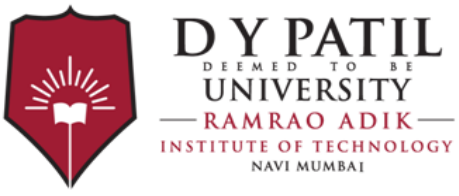
(Computer Engineering)

By

Eshaan Patil	**21IT5040**
Poorva Shetye	**21IT5058**
Ved Bhanushali	**21IT5042**
Rahil Khambe	**21IT5047**
Pratik Bhusare	**21IT5069**

Supervisor

Mrs. Poonam Ghadge



Department of Computer Engineering
Ramrao Adik Institute of Technology,
Sector 7, Nerul, Navi Mumbai
(Under the ambit of D. Y. Patil Deemed to be University)

May 2023

CONTENTS

Chapter 1

Abstract

Speech Recognition is a technology with the help of which a machine can acknowledge the spoken words and phrases, which can further be used to generate text. Speech Recognition System works using techniques popularly termed as acoustic modeling and language modeling. Acoustic modeling represents statistical relationship between the linguistic segments of audio signals and phonemes, on the other hand language modeling represents probability distribution of word segments in a given word sequence.

Today, voice and natural language processing are at the forefront of any human machine interaction environment. The chapter emphasizes the tremendous progress that has taken place in machine learning, statistical data-mining and pattern recognition approaches that can help in making speech interfaces more versatile and pervasive. The growing requirements of speech interfaces also warn against the impediments that may come in the way of successful implementation of acoustically robust natural interface. Finally, the chapter underlines the technical advances and research efforts to be undertaken for high performance realtime speech recognition that will completely change the way humans interact with their computing devices.

Chapter 2

Introduction

- Voice or speaker recognition is the ability of a machine or program to receive and interpret dictation or to understand and perform spoken commands. Voice recognition has gained prominence and use with the rise of artificial intelligence (AI) and intelligent assistants, such as Amazon's Alexa and Apple's Siri.

- Voice recognition systems let consumers interact with technology simply by speaking to it, enabling hands-free requests, reminders and other simple tasks.

- Voice recognnition can identify and distinguish voices using automatic speech recognition (ASR) software programs. Some ASR programs require users first *train* the program to recognize their voice for a more accurate speech-to-text conversion. Voice recognition systems evaluate a voice's frequency, accent and flow of speech.

- Although voice recognition and speech recognition are referred to interchangeably, they aren't the same, and a critical distinction must be made. Voice recognition identifies the speaker, whereas speech recognition evaluates what is said.

- Along with machine learning other technologies which are equally important are IoT, NLP, Big data access management. The use of voice assistants can ease out a lot of tasks for us. Just give voice command input to the system and all tasks will be completed by the assistant starting from converting your speech command to text command then taking out the keywords from the command and execute queries based on those keywords.

Chapter 3

Literature Survey

Sr.No.	Topic	Author	Description
1	Python Based AI Assistant for Computer	Shiv Prakash, Arpit Khare, Sudha Singh,Amisha Gangwar	It's named as Python based AI Assistant for Computer which takes the user input in form of voice or text and process it and returns the output in various forms like action to be performed or the search result is dictated to the end user.
2	Voice assistant using Python	Pooja C. Goutam, Monika S.Jalpure, Akshata S,Gavade, Pranjali Chaudhary, Prof.A.V Gundavade	IT is personal Desktop based voice assistant using Python which is built using open-source software PyCharm as an implementation tool.

3	Speech Recognition	Anjali I.P, Sherseena P. M	Automatic speech recognition is the process by which a computer maps an acoustic speech signal to text. Automatic speech understanding is the process by which a computer maps an acoustic speech signal to some form of abstract meaning of the speech.
4	Voice Assistant using Artificial Intelligence	Ms. Preethi G, Mr. Abishek K, Mr. Thiruppugal S, Mr. Vishwaa D A	A voice assistant can be a digital assistant that uses human voice, language process algorithms, and synthesis to pay attention to particular voice commands and come applicable information or perform particular functions as appealed by the user supported commands

Chapter 4

Problem Statement

We are all well aware about Cortana, Siri, Google Assistant and many other virtual assistants which are designed to aid the tasks of users in Windows, Android and iOS platforms. But to our surprise, there's no such complete virtual assistant available for Core Windows platform consisting of 70% of the users. So, this is actually a major problem for users where there could be internet instability, server problems and places where internet is not accessible.

Chapter 5

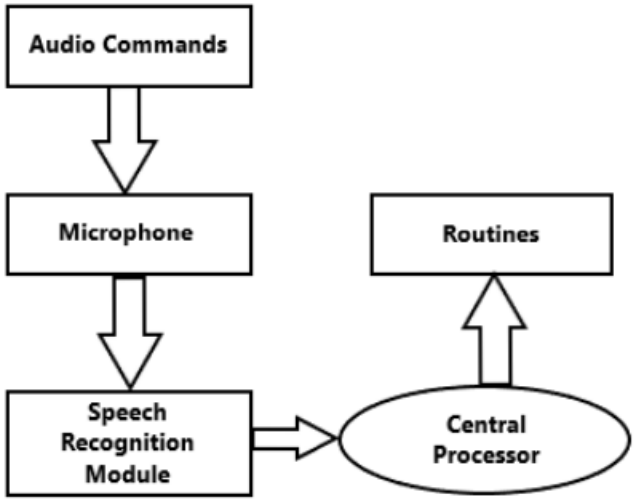
Proposed System

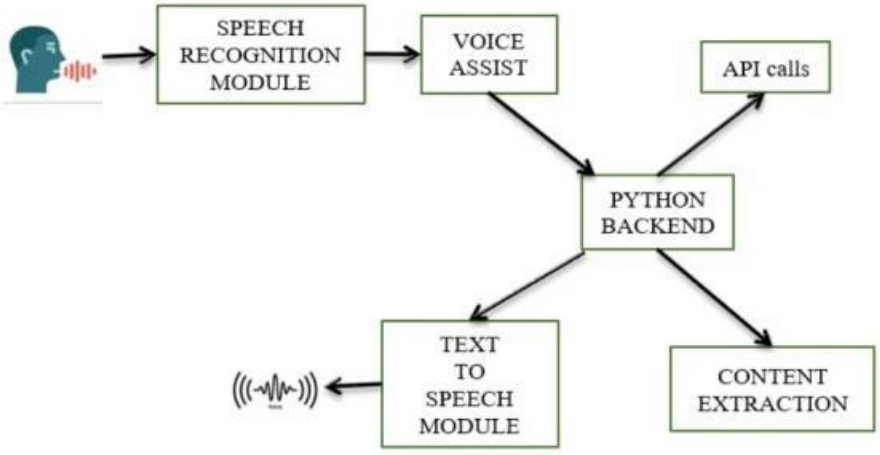
In this proposed concept effective way of implementing a Personal voice assistant, Speech Recognition library has many in-built functions, that will let the assistant understand the command given by user and the response will be sent back to user in voice, with Text to Speech functions. When assistant captures the voice command given by user, the under lying algorithms will convert the voice into text.

Proposed Architecture

The system design consists of

1. Taking the input as speech patterns through microphone.
2. Audio data recognition and conversion into text.
3. Comparing the input with predefined commands.
4. Giving the desired output.





Chapter 6
Methodology

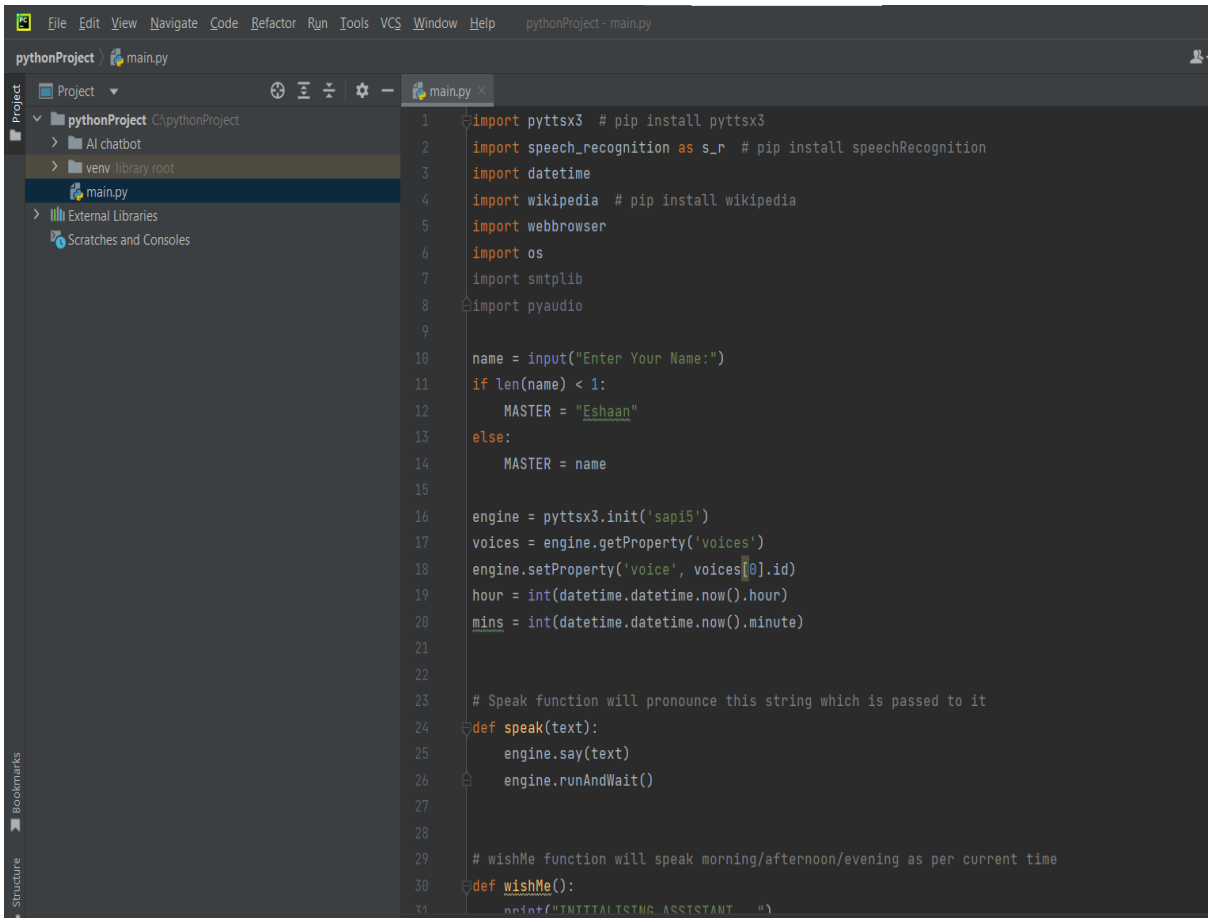
Voice assistant applications work based on Automatic Speech Recognition (ASR) system. ASR systems record the speech and then break it down into phonemes, which are later get processed into text. A phoneme (not words of syllables) is a basic unit of measurement for human speech recognition**.**

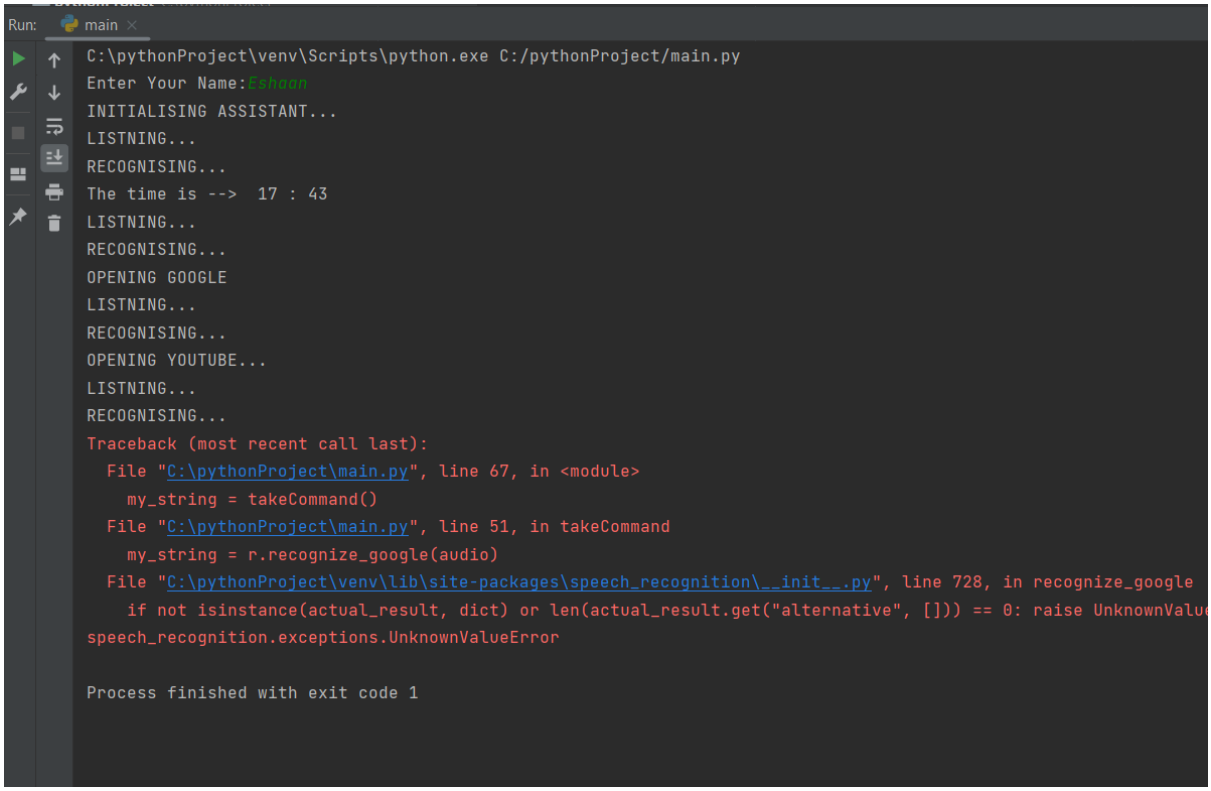
Phoneme recognition delivers better results than the process of word decoding, as the last one tends to analyze word as a standalone unit ignoring the context.

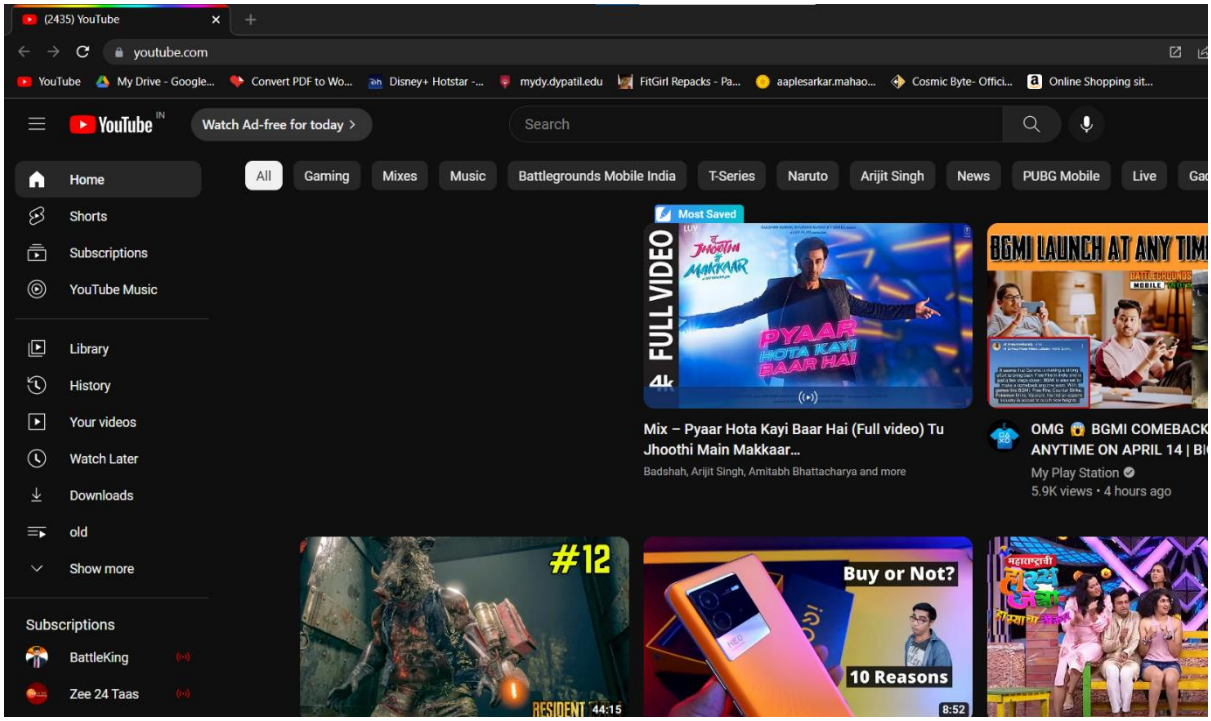
Acoustic modeling, which represents the which phonemes were pronounced and what are the words these phonemes complete; Pronunciation modeling, that analyzes the way phonemes are pronounced, is there any accent or other peculiarities of the vocal apparatus to capture the phonetic variability of speech; Language modeling, which is aimed at finding contextual probabilities depending on what phonemes were captured.

Chapter 7

Results




Run:
main
C:\pythonProject\venv\Scripts\python.exe C:/pythonProject/main.py
Enter Your Name:Eshaan
INITIALISING ASSISTANT...
LISTNING...
RECOGNISING...
The time is --> 17 : 43
LISTNING...
RECOGNISING...
OPENING GOOGLE
LISTNING...
RECOGNISING...
OPENING YOUTUBE...
LISTNING...
RECOGNISING...
Traceback (most recent call last):
File "C:\pythonProject\main.py", line 67, in <module>
my_string = takeCommand()
File "C:\pythonProject\main.py", line 51, in takeCommand
my_string = r.recognize_google(audio)
File "C:\pythonProject\venv\lib\site-packages\speech_recognition__init__.py", line 728, in recognize_google
if not isinstance(actual_result, dict) or len(actual_result.get("alternative", [])) == 0: raise UnknownValue
speech_recognition.exceptions.UnknownValueError
Process finished with exit code 1


(2435) YouTube
youtube.com
YouTube
Watch Ad-free for today
Search
Home
Shorts
Subscriptions
YouTube Music
Library
History
Your videos
Watch Later
Downloads
old
Show more
Subscriptions
BattleKing
Zee 24 Taas
All
Gaming
Mixes
Music
Battlegrounds Mobile India
T-Series
Naruto
Arijit Singh
News
PUBG Mobile
Live
Most Saved
FULL VIDEO
PYAAR HOTA KAYI BAAR HAI
Mix – Pyaar Hota Kayi Baar Hai (Full video) Tu Jhoothi Main Makkaar...
Badshah, Arijit Singh, Amitabh Bhattacharya and more
BGMI LAUNCH AT ANY TIM
OMG BGMI COMEBACK ANYTIME ON APRIL 14 | BI
My Play Station
5.9K views • 4 hours ago
#12
RESIDENT
44:15
Buy or Not?
10 Reasons
8:52

Chapter 8

Conclusion

- An excellent virtual assistant will save time and money by doing the small tasks for you and doing them accurately and with high quality. If you handle the virtual assistant correctly, it will be a boom in your business. If you'd like to find out more about hiring a virtual assistant, please consider someone from VP Virtual Assistants.
- The future of voice search and assistants is looking bright. With the number of people already seeing how convenient those tools can be and the growing number of devices that use Voice Recoginition. It's clear that the technology will soon be everywhere, and with 5G and improvements in machine learning, voice assistants might at some point become tools we can't live without.
- Through this voice assistant, we have automated various services using a single line command. It eases most of the tasks of the user like searching the web, opening application etc