

**SUMMER INTERNSHIP PROJECT**

**Submitted By**

**Name /Semester/Course/Enrollment No. –**

ANSHUL SINHA / III / BCA / A71004819044

PRATHMESH PATIL / III / BCA / A71004819033

NINAD NIRBHAVANE/ III / BCA / A71004819035

**Guided By-**

Professor Kirti Kakde

**AMITY UNIVERSITY MAHARASHTRA**

**Amity Institute of Information Technology**

Amity Institute of Information Technology



Certificate

Certified that the summer internship project work entitled “AI VOICE ASSISTANT” carried out by ANSHUL SINHA(A71004819044), PRATHMESH PATIL(A71004819033) and NINAD NIRBHAVANE(A71004819035) from BCA Batch are Bonafide students of Amity Institute of Information Technology and have successfully completed the Summer Internship Project of 3rd Semester in the academic year 2020.

Subject Teacher’s signature

**INDEX**

1) PROJECT TITLE

2) INTRODUCTION

3) AIM AND PURPOSE

4) TECHNOLOGY USED

5) MODULE DESCRIPTION

6) VOICE ASSISTANT ON OUR SYSTEM

7) PROPOSED SYSTEM

8) HISTORY OF VOICE ASSISTANT

9) UML DIAGRAMS

10) SCREENSHOT OF OUTPUT WITH VALIDATION

11) RESULT AND DISCUSSION

12) FUTURE SCOPE

13) REFERENCE AND BIBLIOGRAPHY

**PROJECT TITLE**

“AI VOICE ASSISTANT”

**INTRODUCTION**

This project is based on desktop application development and provide personal assistant using voice recognition, with the help of python programming language. This program includes the functions and services of: find location, convert currencies of different countries, tell current time, can tell you the latest news, can calculate basic mathematical calculations, can answer your any general knowledge question, mail exchange, event handler, music player service, can tell you jokes, checking weather, Google searching engine, Wikipedia searching engine, direct access and search in popular e- commerce sites like amazon and flipkart, can also do direct search on entertainment platforms like Netflix, Amazon Prime Video and YouTube and initiate the result, can open almost all of the Microsoft office applications like Microsoft PowerPoint, Excel or Word and there are many other tasks which the voice assistant can do just by the voice command of the user.

Wake words rely on a special algorithm that is always listening for a particular word or phrase so that a phone, smart speaker, or something else can begin communicating with a server to do its job. Wake words need to be long enough to be distinct, easy for a human to speak, and simple for a machine to recognize. This is why we cannot change our wake word to anything we want it to be.  Voice assistants don’t really “understand” what we are saying — they just listen for their wake word and then begin communicating with a server to complete a task.

As it integrates most of the desktop services for daily use, it could be useful for getting a more convenient life and it will be helpful for those people who have disabilities for manual operations.

**AIM AND PURPOSE**

The purpose of the project is to develop a Desktop application that provides an intelligent voice assistant with sorts of different virtual functionalities. Nowadays, computers are commonly and widely used by all people. It is not a convenient way for users with completely manually input. The common way of communication used by people in daily life is through the speech. If the system can just listen to the user for the request or handle the daily affairs, then give the right response, it will be much easier for users to communicate with their system, and it will be much “Smarter” as a human assistant. This project is focusing on the voice control (recognition, generate and analyze corresponding commands, intelligent responses automatically), Google products and relevant APIs (Google map, Google weather, Google search and etc.), Wikipedia API and desktop references. As all those functionalities and services for the project have been explained, the main structure and construction of the project has been basically illustrated with its goals.

**TECHNOLOGY USED**

SOFTWARE USED: -

* VISUAL STUDIO CODE (for Python Programming)
* TKINTER (for graphical interface used in vs code)
* WINDOWS 10 HOME

HARDWARE USED: -

* Intel® Core™ i5-8300H CPU @ 2.30GHz PROCESSOR
* 8.00 GB RAM
* 1 TB HDD STORAGE CAPACITY

**MODULE DESCRIPTION**

This project is all about the creation of an AI voice Assistant which is made by using Python programming language.

We have worked with python in Visual Studio Code, using the Microsoft Python extension.

We have defined various functions like- Speak Function (this function will program our voice assistant to speak something), Take Command Function (this function will allow our voice assistant to take microphone input from the user and returns a string output). Wish me Function (this function will make our voice assistant to wish us according to system time).

To make our voice assistant perform majority of different functions, we have imported a lot of modules from python which includes –

\*pyttsx3

\* webbrowser

\*os

\*Wikipedia

\*datetime

\*speech recognition as sr

\*smtplib,

\*random

\*youtube\_dl

\*pyjokes

\*wolframalpha

\*subprocess.

This program includes the functions and services of: -

Find location, convert currencies of different countries, can tell you the latest news, can calculate basic mathematical calculations, can answer your any general knowledge question, mail exchange, tell current time, event handler, music player service, can tell you jokes, checking weather, Google searching engine, Wikipedia searching engine, direct access and search in popular e- commerce sites like amazon and flipkart, can also do direct search on entertainment platforms like Netflix, Amazon Prime Video and YouTube and initiate the result, can open almost all of the Microsoft office applications like Microsoft PowerPoint, Excel or Word and there are many other tasks which the voice assistant can do just by the voice command of the user.

The next step taken was to give the code a basic graphical interface. So, we use Tkinter/ TK interface for fulfilling the interface part.

Tkinter is the standard GUI library for Python. Python when combined with Tkinter provides a fast and easy way to create GUI applications. Tkinter provides a powerful object-oriented interface to the Tk GUI toolkit

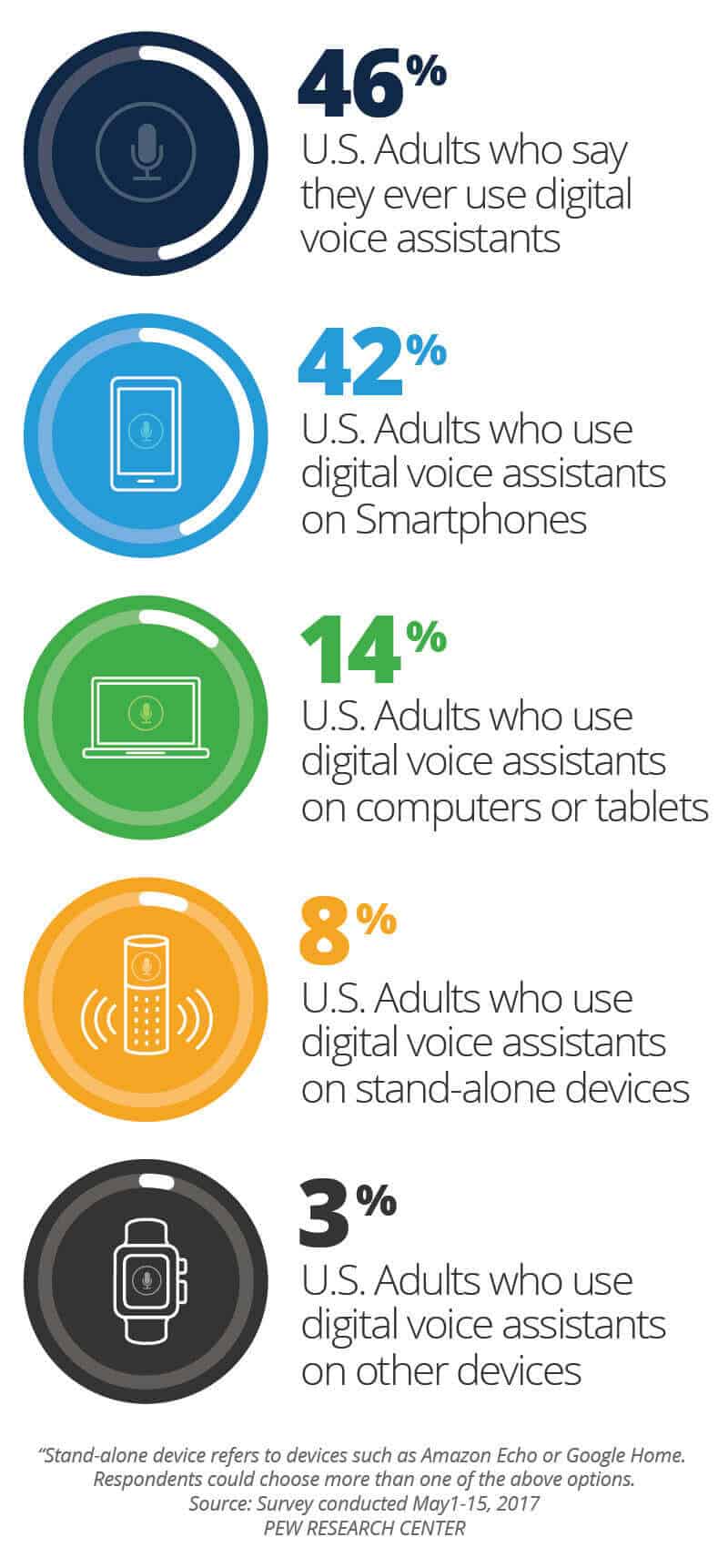
This project is focusing on the desktop assistant over the voice control (recognition, generate and analyze corresponding commands, intelligent responses automatically), to make our interaction with the computer more easily and efficiently.

**Voice Assistants on Our System**

Voice assistants allow us to do a variety of tasks hands-free, which is a major reason many people like using them. Apple has Siri. Google phones and most Androids have Google. Samsung has Bixby. Windows phones have Cortana.  
A Pew Research Center survey in May 2017 showed that nearly half of all adults in the United States use voice-controlled digital assistants on their devices.

Voice assistants can make calls, send text messages, look things up online, provide directions, open apps, and initiate or complete many other tasks.

With the addition of separate apps on the device, our voice can be a type of remote control for our lives.



**PROPOSED SYSTEM**

The proposed system will provide following features:

* It always keeps listing for its name and wakes up to response upon calling with the assigned functionality.
* Performing Wikipedia search based on voice commands and giving back the computed response through a voice.
* Searching Internet based on user’s voice input and giving back the reply through a voice with further interactive questions by machine.
* Direct access to e-commerce sites like amazon and flipkart and can give you the products search result through user’s voice. User just have to speak the name of the item they want to buy/view and the voice assistant will give the list of that item.
* Direct access to entertainment sites like Netflix, Amazon Prime Video and YouTube and can give you the search result of that particular show the user wants to watch.
* Other features such as-: find location, convert currencies of different countries, can tell you the latest news, can calculate basic mathematical calculations, tell current time, can answer your any general knowledge question, mail exchange, event handler, music player service, can tell you jokes, checking weather, Performing Wikipedia search, Google searching engine, can open almost all of the Microsoft office applications like Microsoft PowerPoint, Excel or Word and there are many other tasks which the voice assistant can do just by the voice command of the user.

**The History of Voice Assistants**



Technology companies are working to create increasingly sophisticated technology that will automate more processes and tasks we do throughout the day. Even Cortana, Siri, Google Assistant, and Alexa can “learn” new words and tasks.

**UML DIAGRAMS**

A UML diagram is a diagram based on the UML (Unified Modelling Language) with the purpose of visually representing a system along with its main actors, roles, actions, artifacts or classes, in order to better understand, alter, maintain, or document information about the system.

In this project, we have shown three different types of uml diagrams for our voice assistant that includes-

1. Activity Diagram
2. Use Case Diagram
3. Sequence Diagram

**ACTIVITY DIAGRAM**

**Activity diagram** is basically a flowchart to represent the flow from one activity to another activity. The activity can be described as an operation of the system. The basic purpose of activity diagrams is to capture the dynamic behaviour of the system.

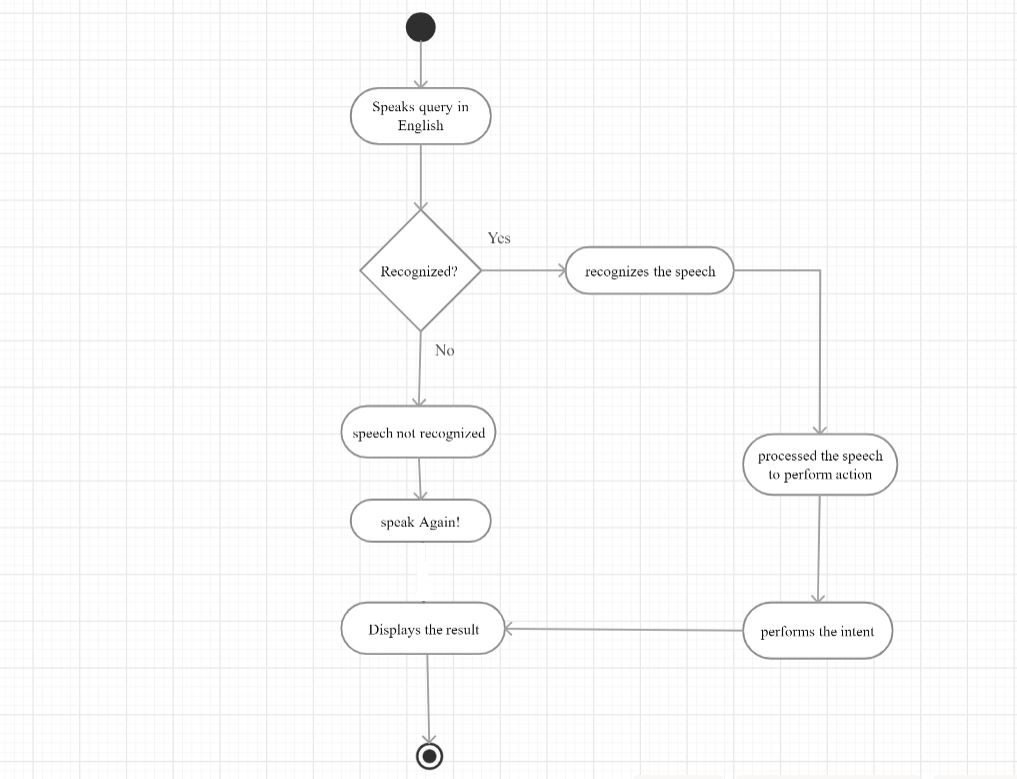
This UML diagram focuses on the execution and flow of the behaviour of a system instead of implementation. Activity diagrams consist of activities that are made up of actions that apply to behavioural modelling technology.

Explanation: -

In the activity diagram given below, the speech recognition activity is specified. When the speech recognition process begins, the user first speaks query in English language then the decision box determines whether the speech is recognized or not.

If the speech is recognized, then it processes the speech to perform the action, it performs the intent and then displays the result.

If the speech is not recognized, then it will ask the user to speak Again.



\* ACTIVITY DIAGRAM

**USE CASE DIAGRAM**

Use case diagram is an uml diagram which can summarize the details of our system's users (also known as actors) and their interactions with the system. To build one, you'll use a set of specialized symbols and connectors. An effective use case diagram can help your team discuss and represent:

* Scenarios in which your system or application interacts with people, organizations, or external systems
* Goals that your system or application helps those entities (known as actors) achieve
* The scope of your system

Explanation: -

In the use case diagram given below, the User(actor) interacts with the system through the input voice for performing the various functions like search google, Wikipedia search, joke, location, send email, search Netflix, search amazon or flipkart.

So, through this interaction a particular goal/work of user can be accomplished.



\* USE CASE DIAGRAM

**SEQUENCE DIAGRAM**

A sequence diagram simply depicts interaction between objects in a sequential order i.e. the order in which these interactions take place. We can also use the terms event diagrams or event scenarios to refer to a sequence diagram.

Sequence diagrams describe how and in what order the objects in a system function. These diagrams are widely used by businessmen and software developers to document and understand requirements for new and existing systems.

Explanation: -

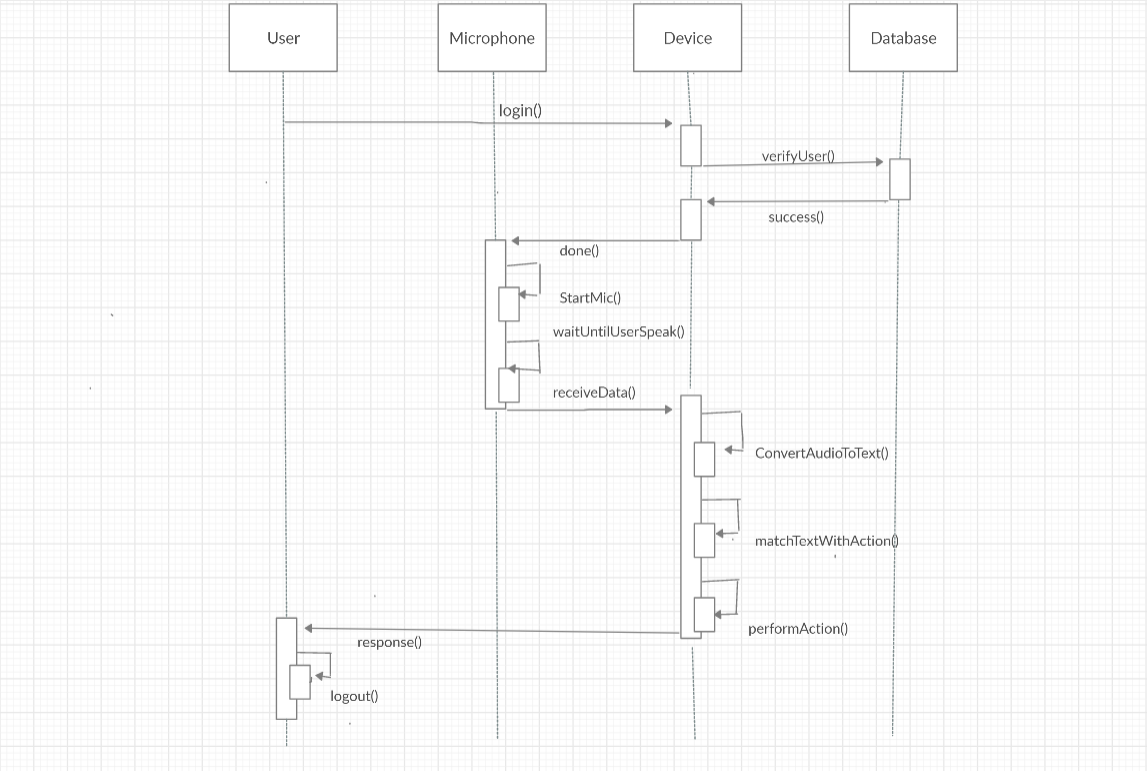
In the sequence diagram given below, there are four objects- User, Microphone, Device and Database.

The user login the device and then verifies through database.

After successful login, the message will be sent to device and then to the microphone. The microphone will get started and then will wait until the user speak any query.

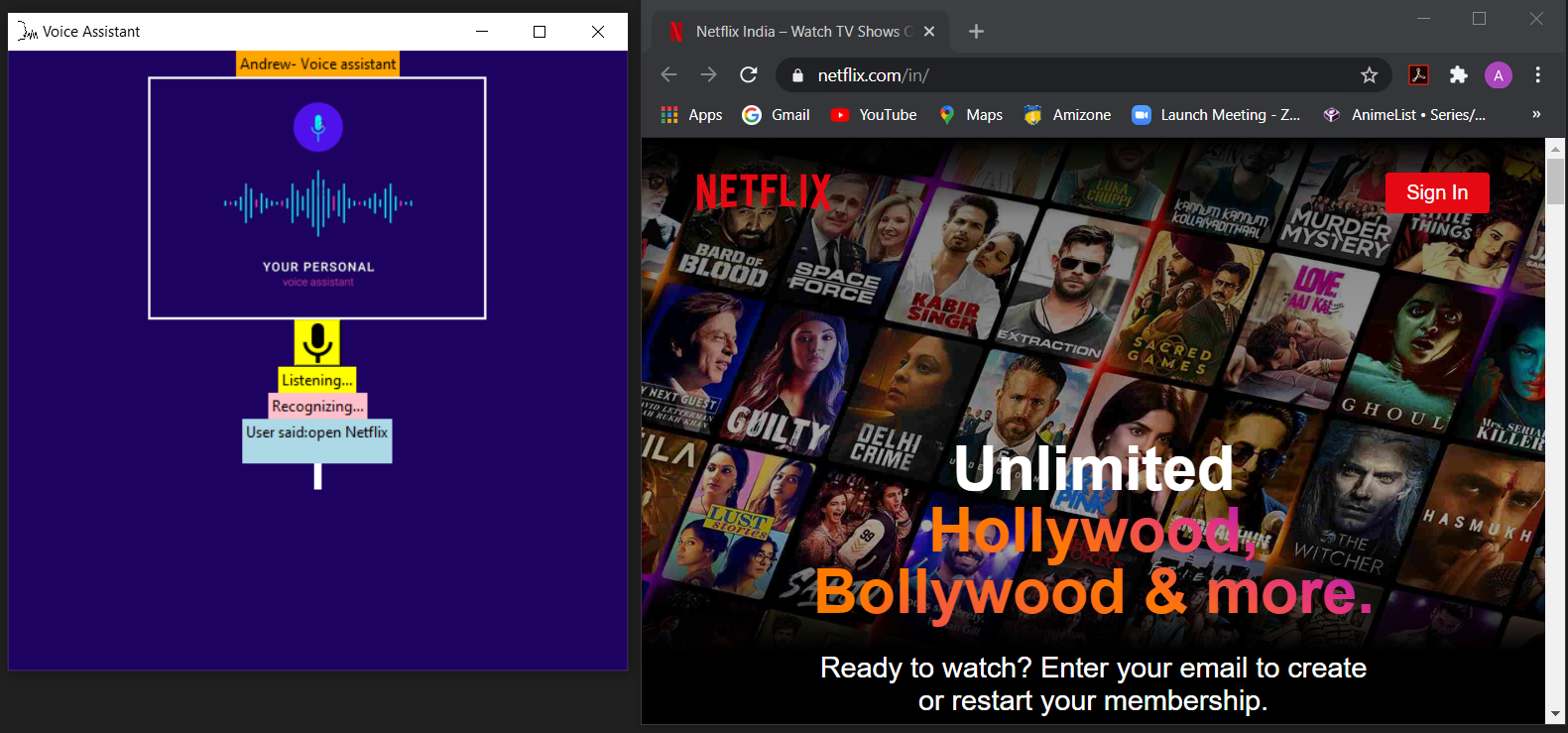
Next the message will be sent to the device and in the object device, the audio will be converted to text. The text will then will be matched with action and will perform the action.

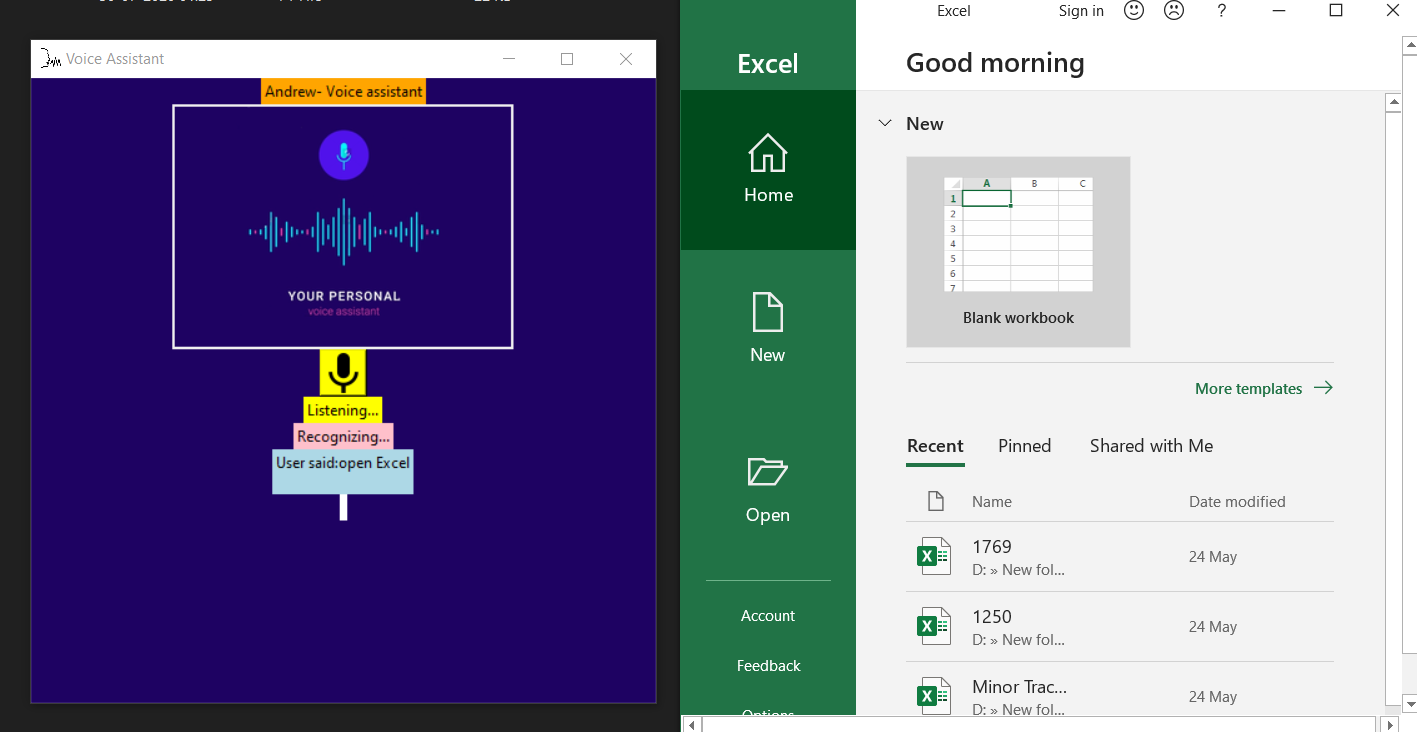
The response message will be sent to the user and after getting the response, the user can logout.

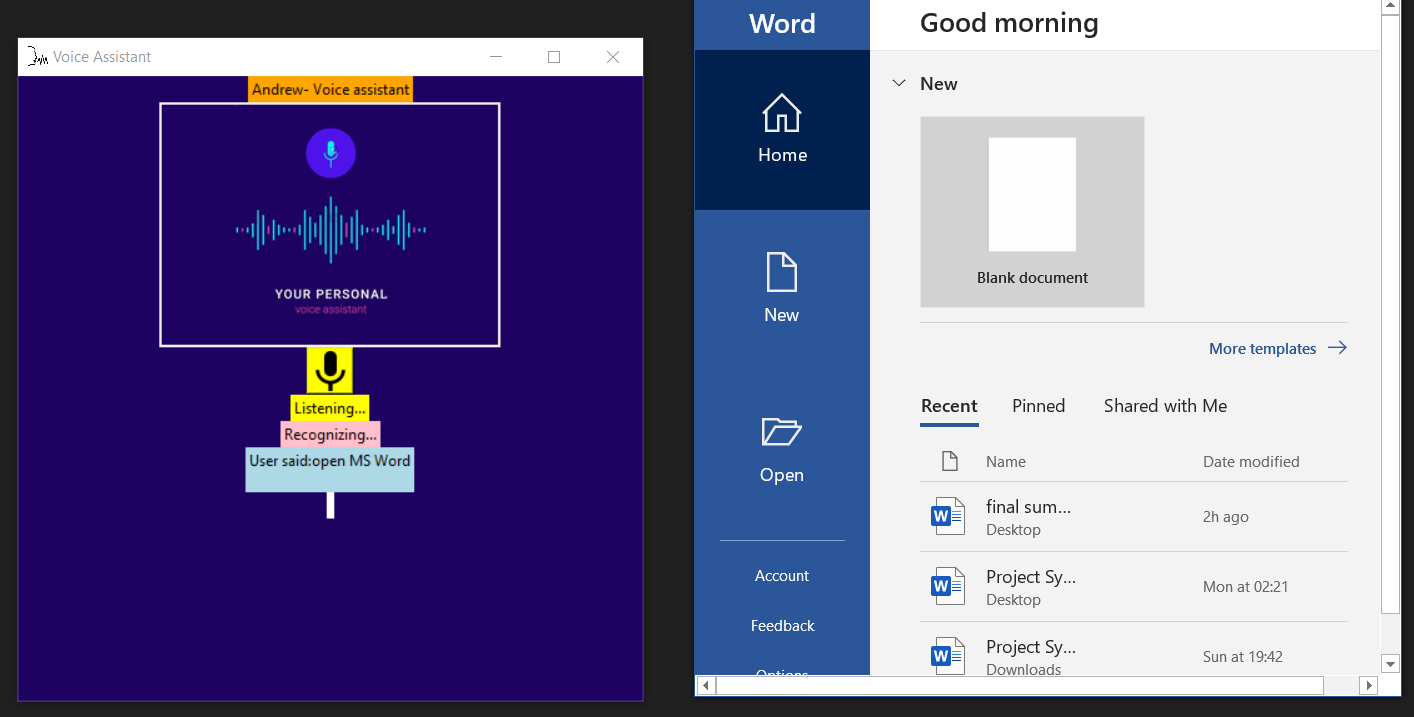


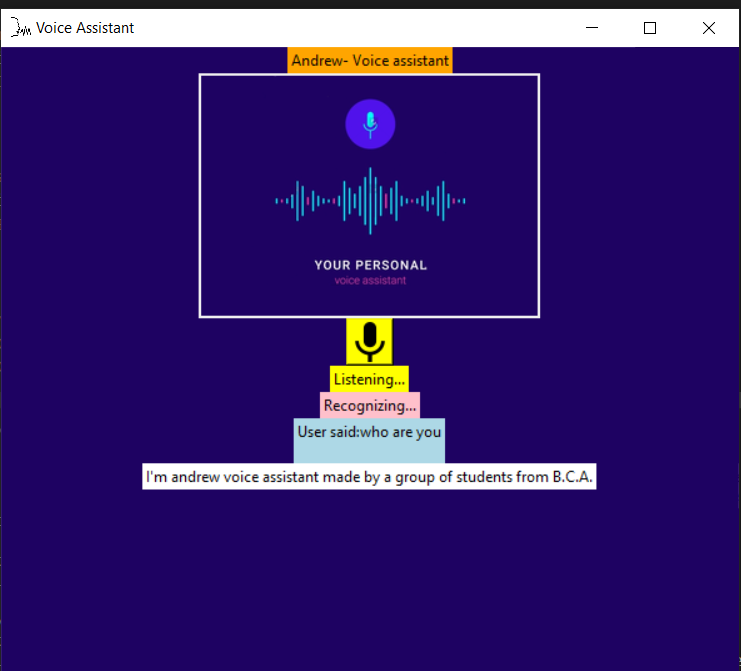
\*SEQUENCE DIAGRAM

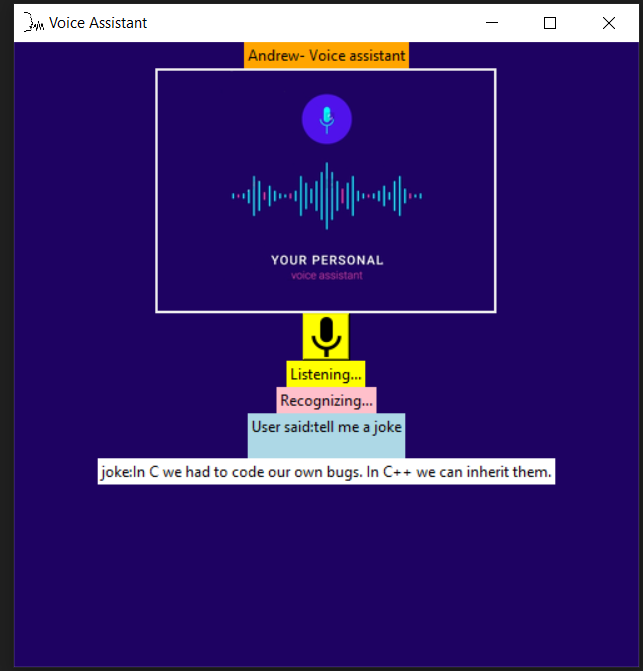
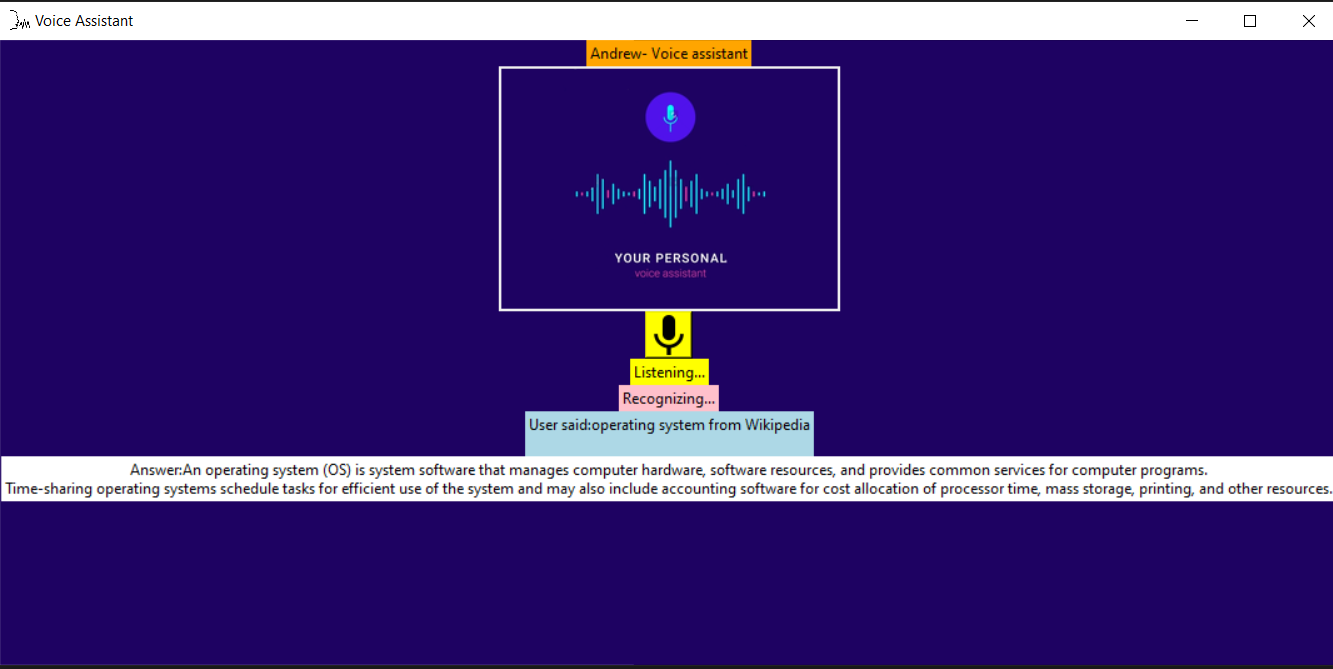
**SCREENSHOT OF OUTPUT WITH VALIDATION**



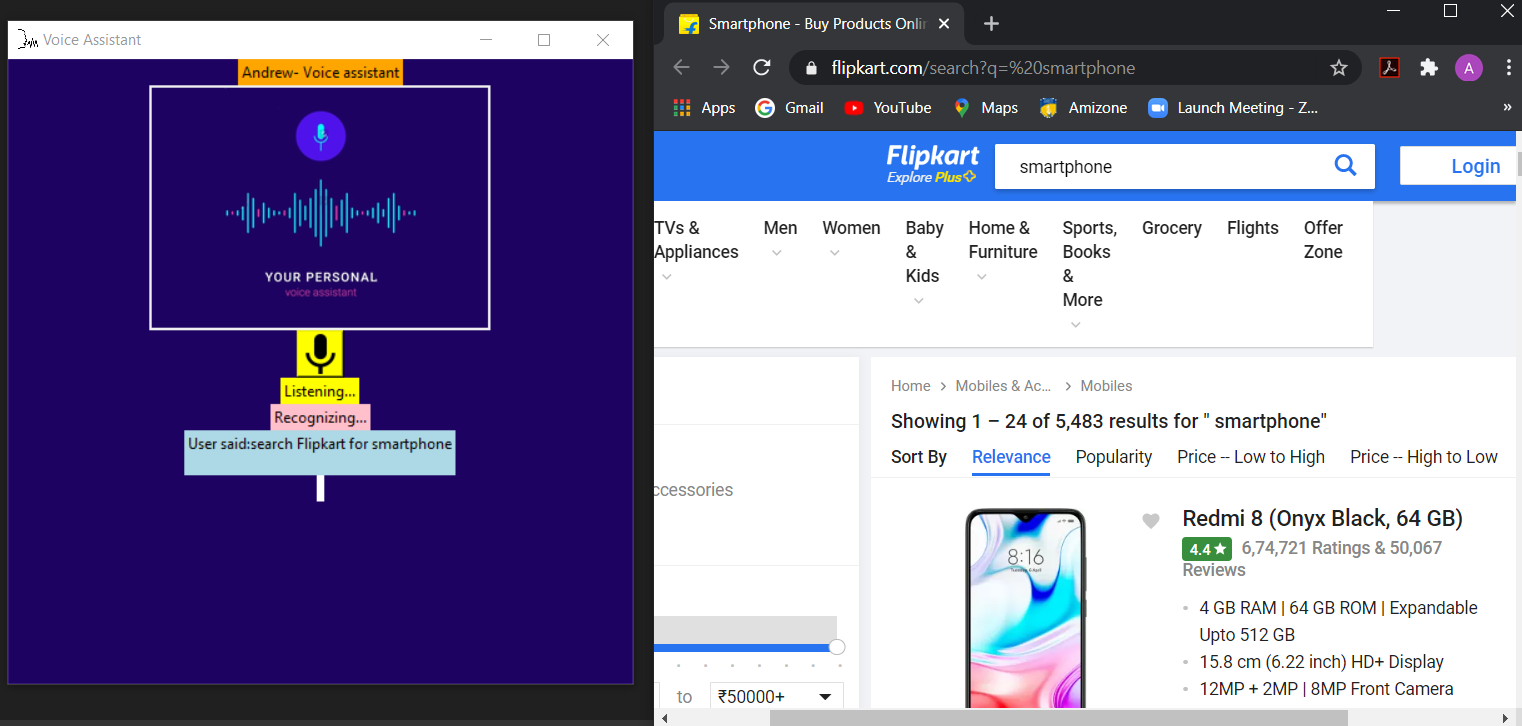


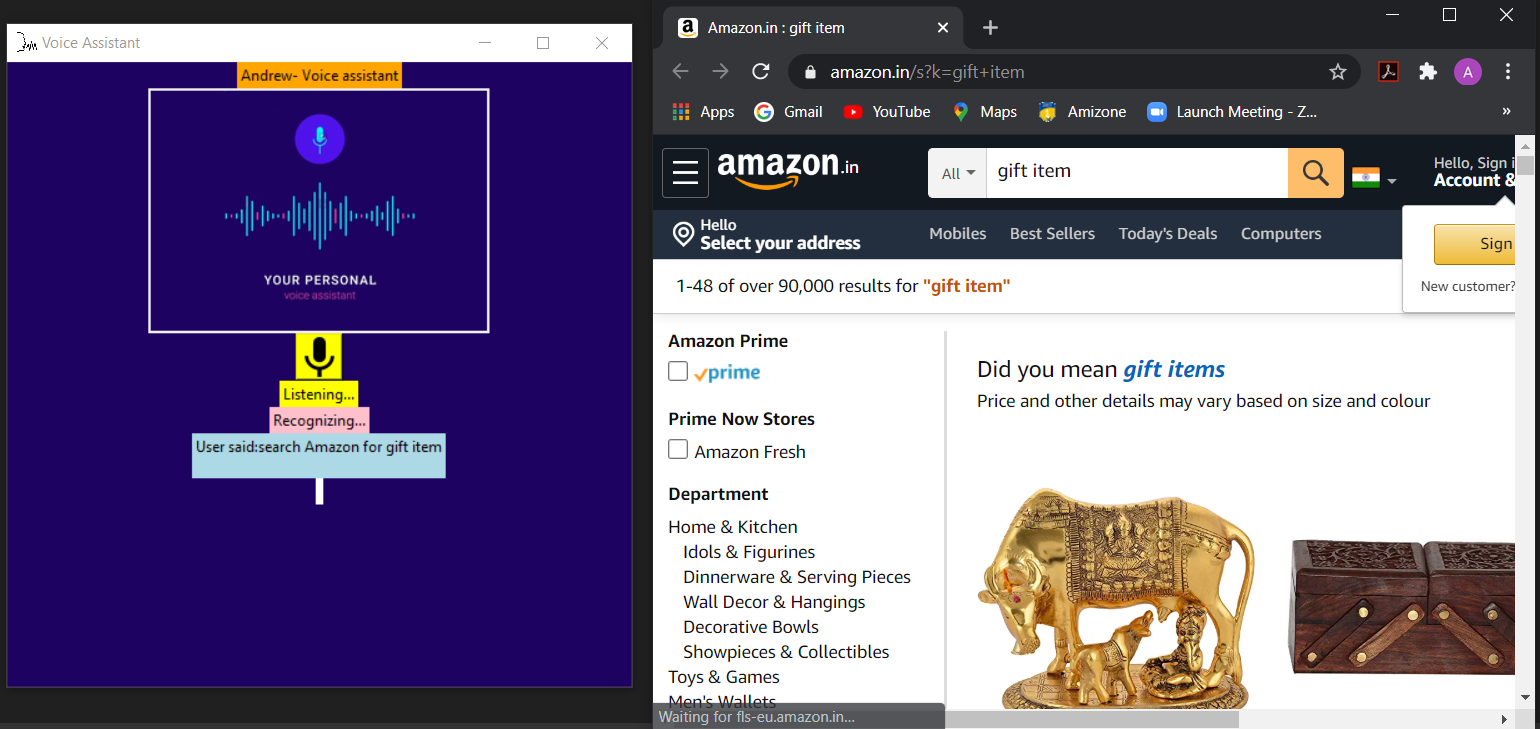


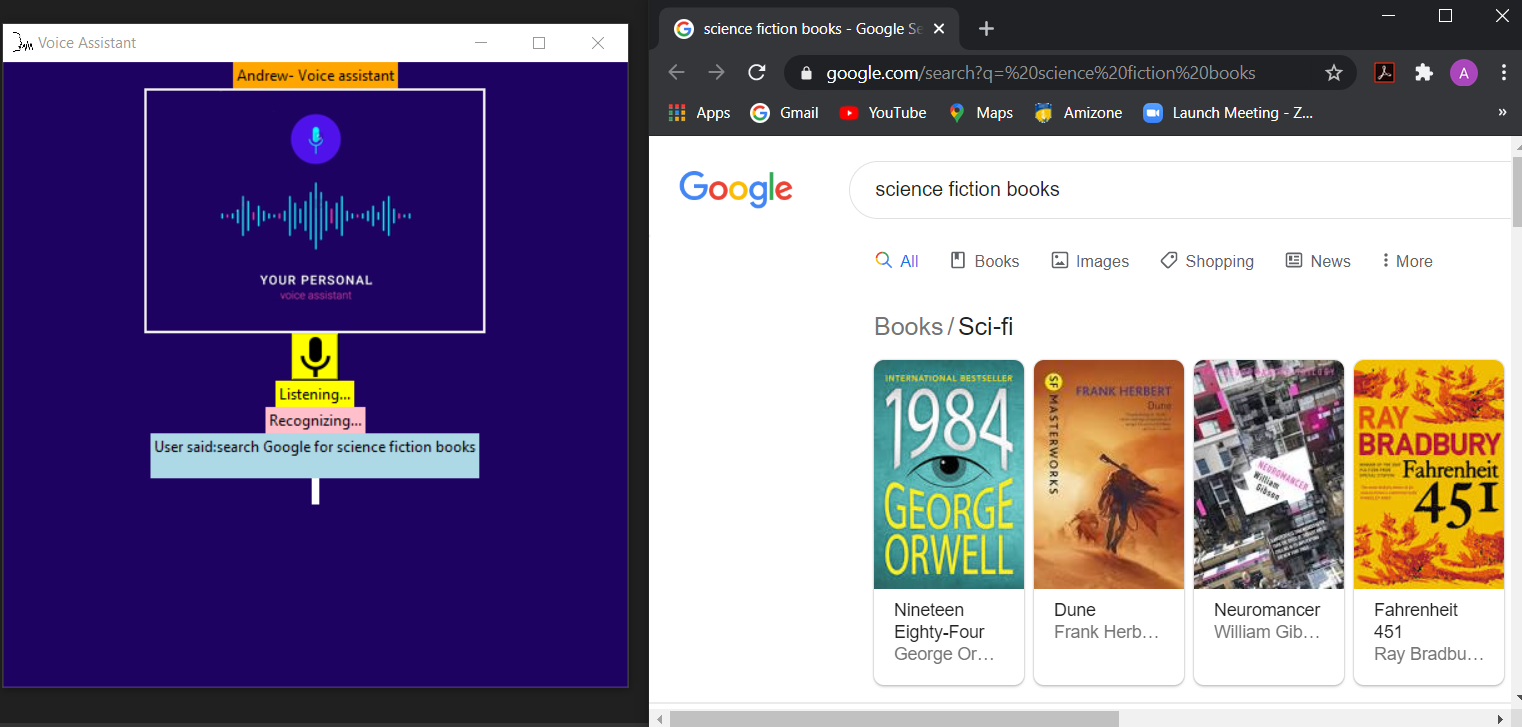


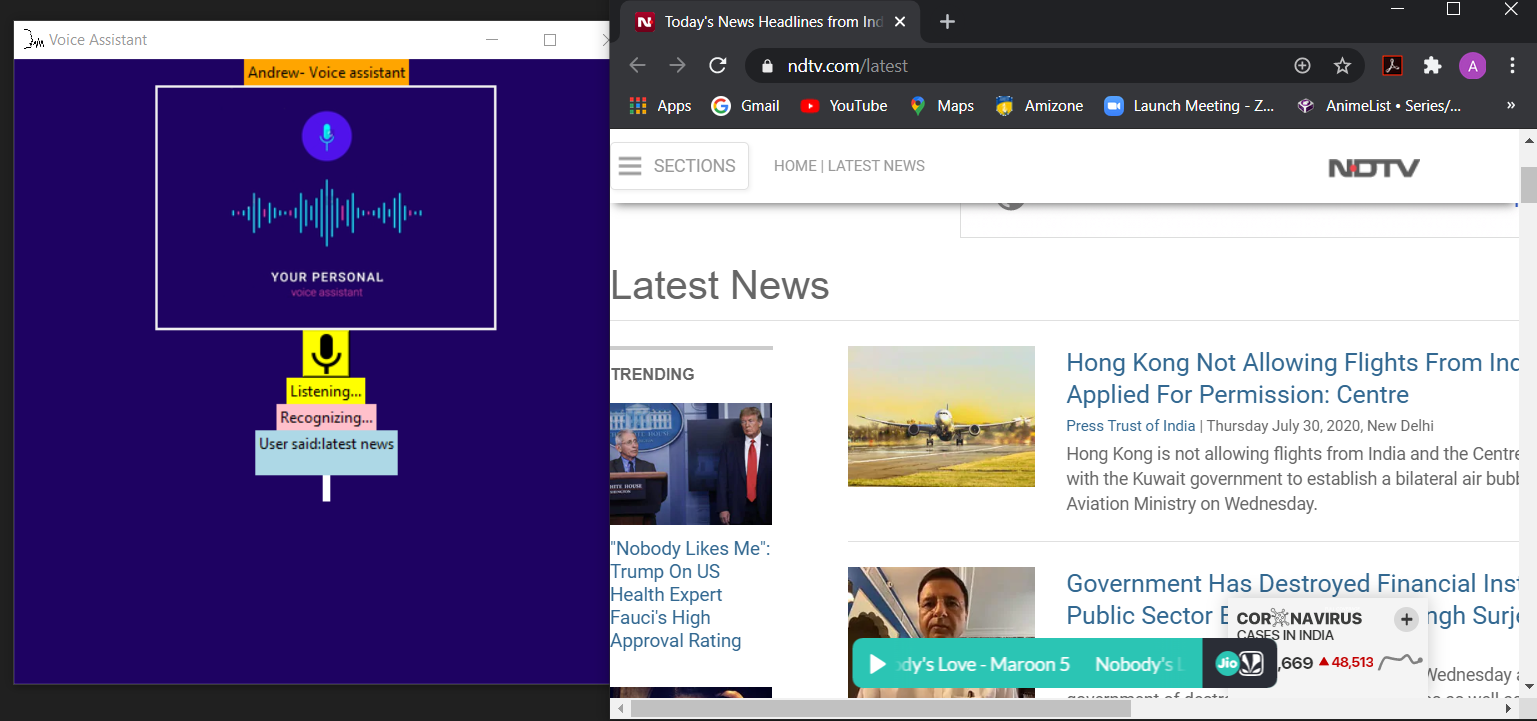


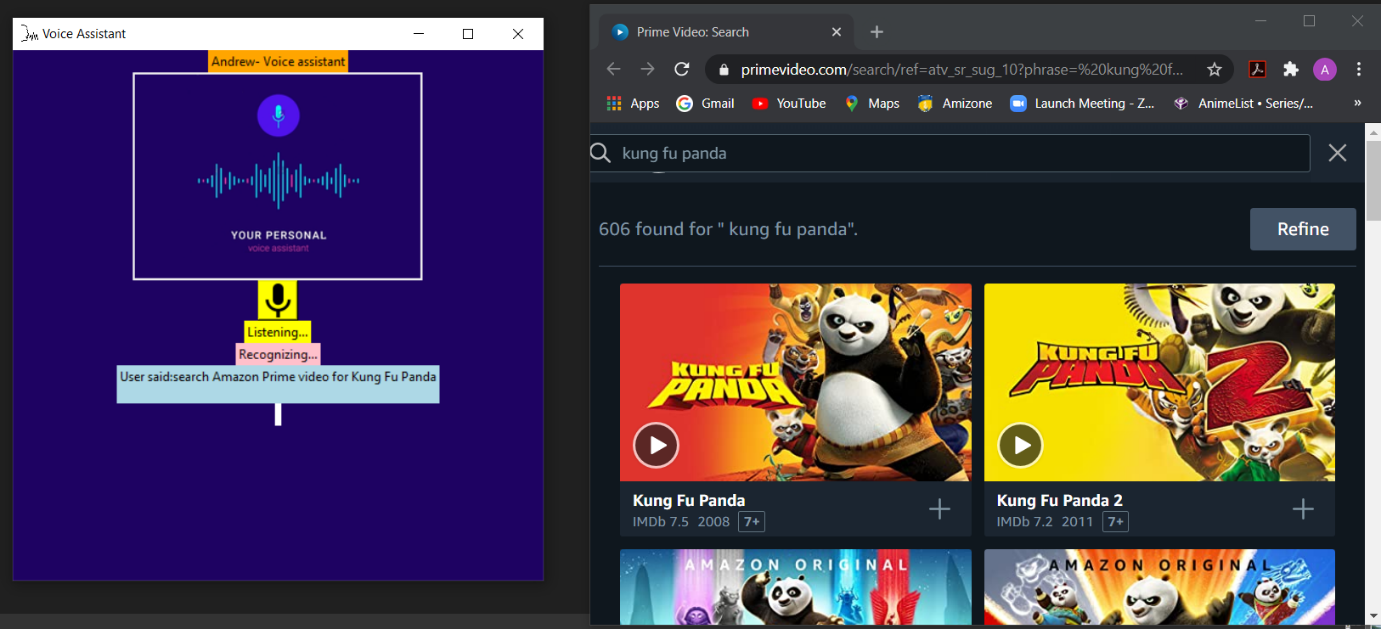
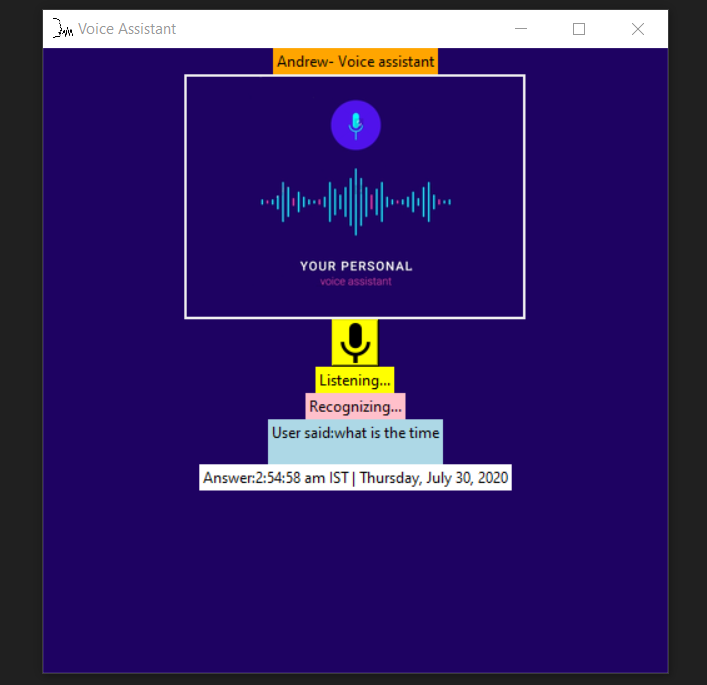
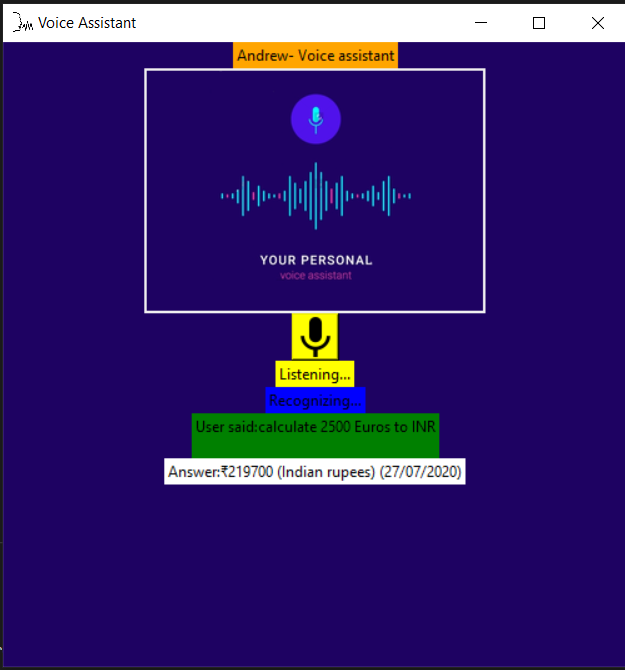
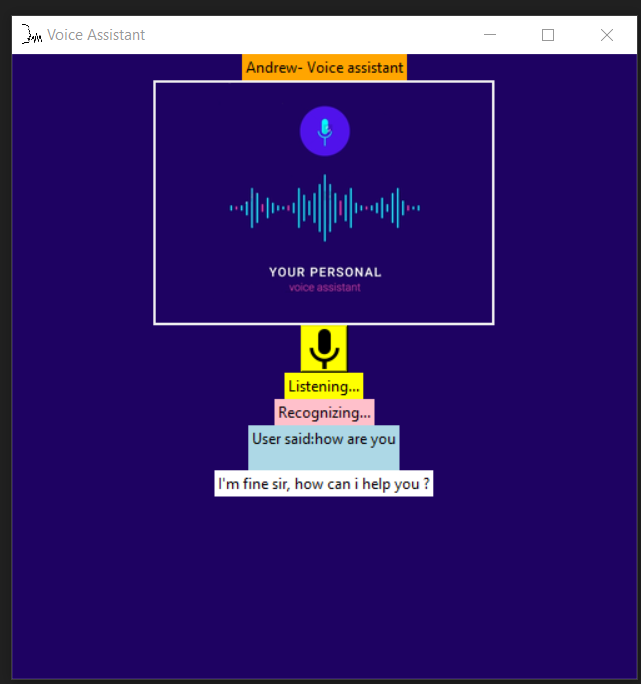
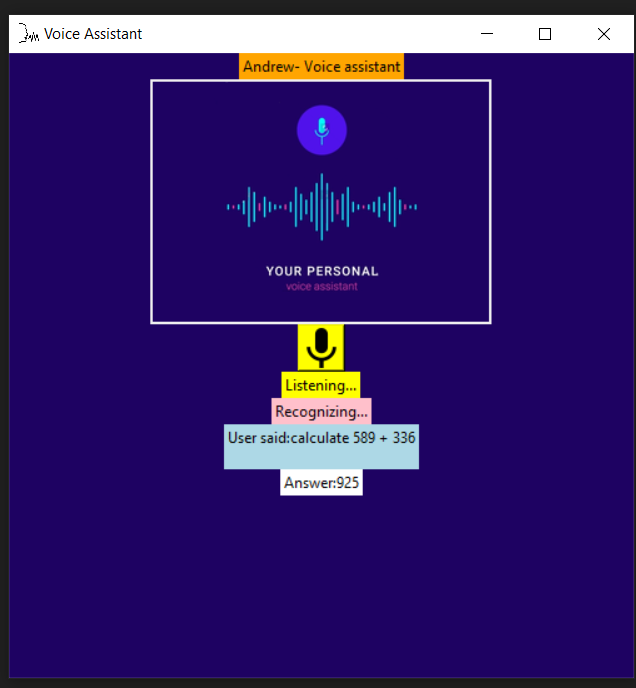
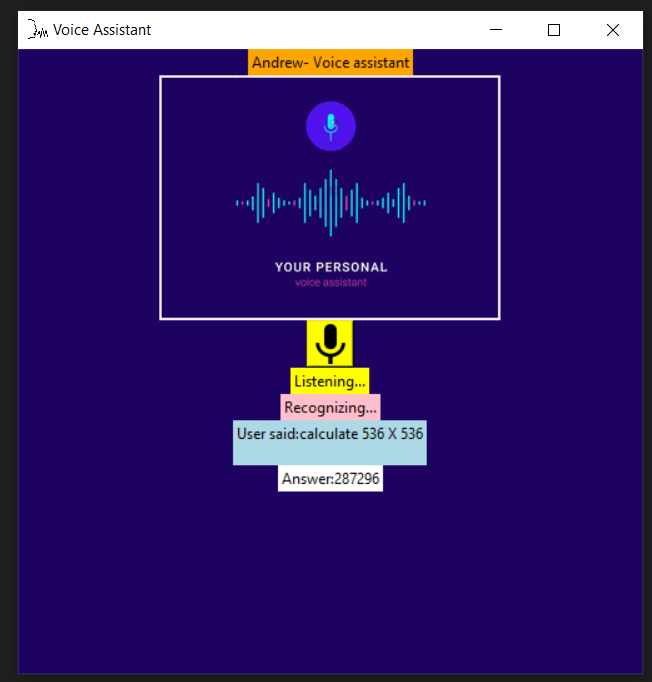
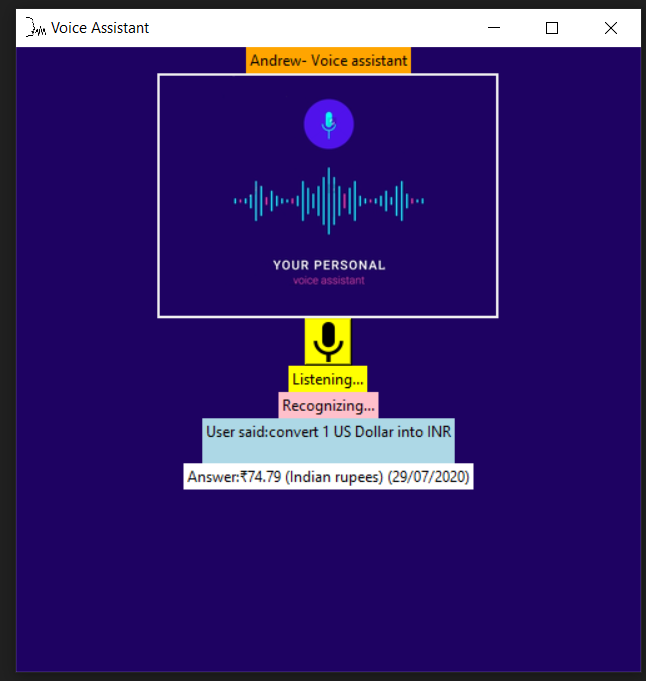
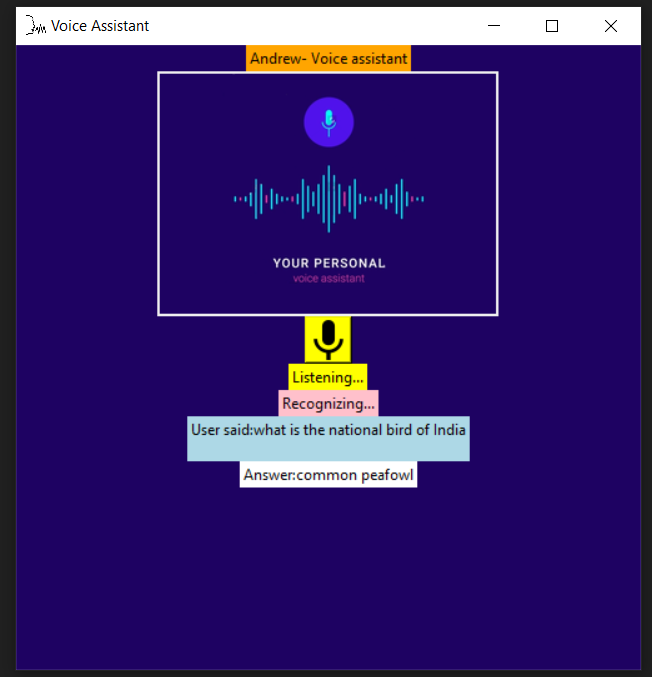
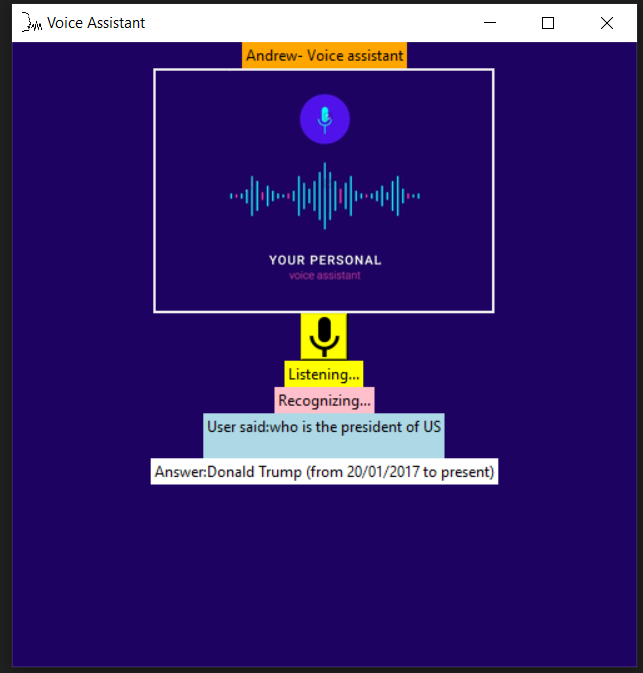
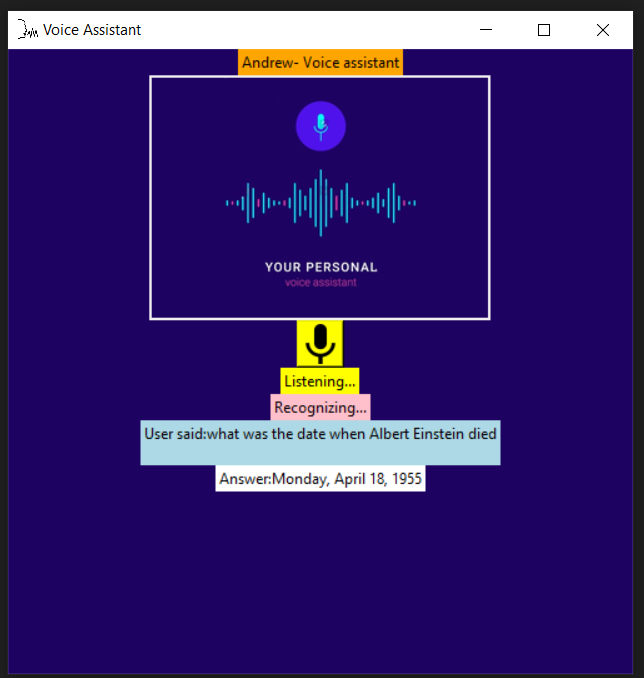
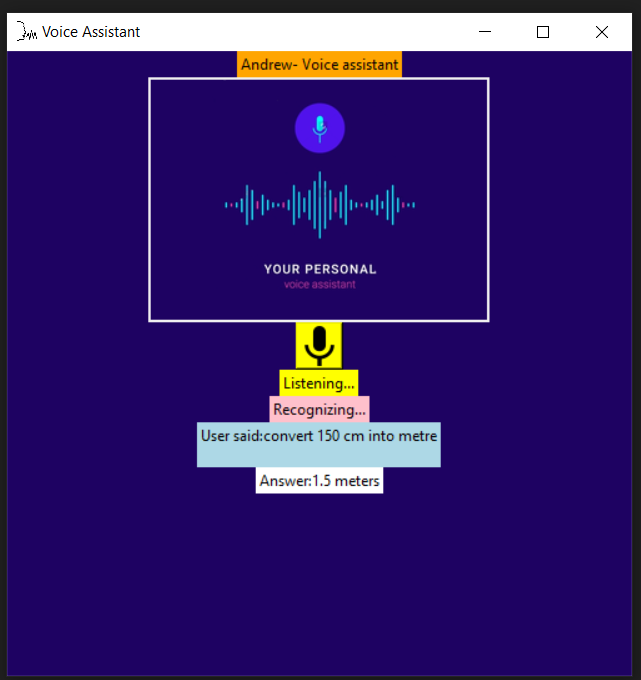
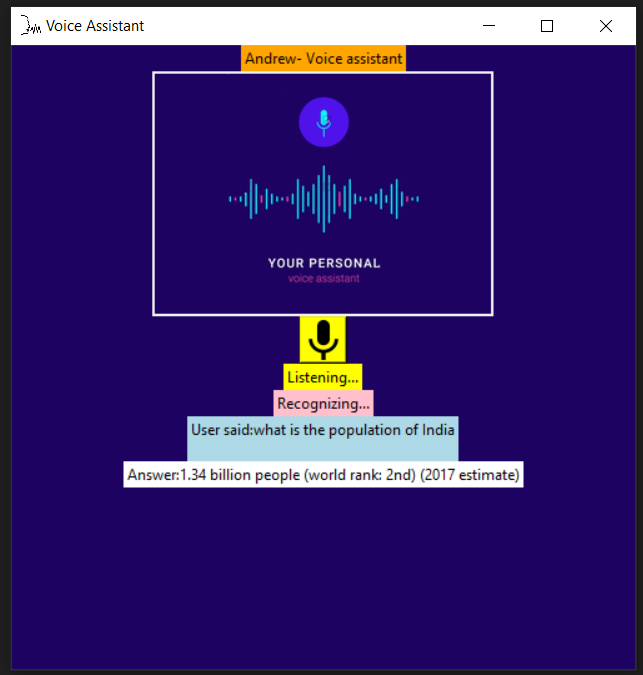
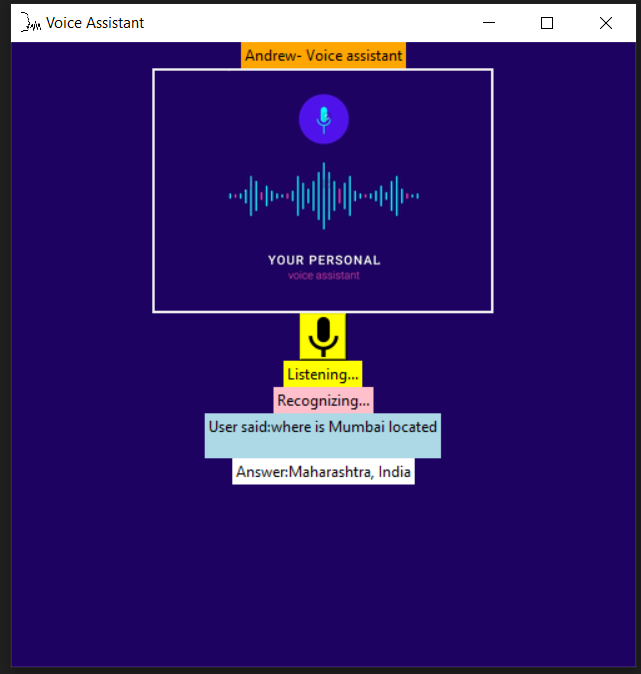
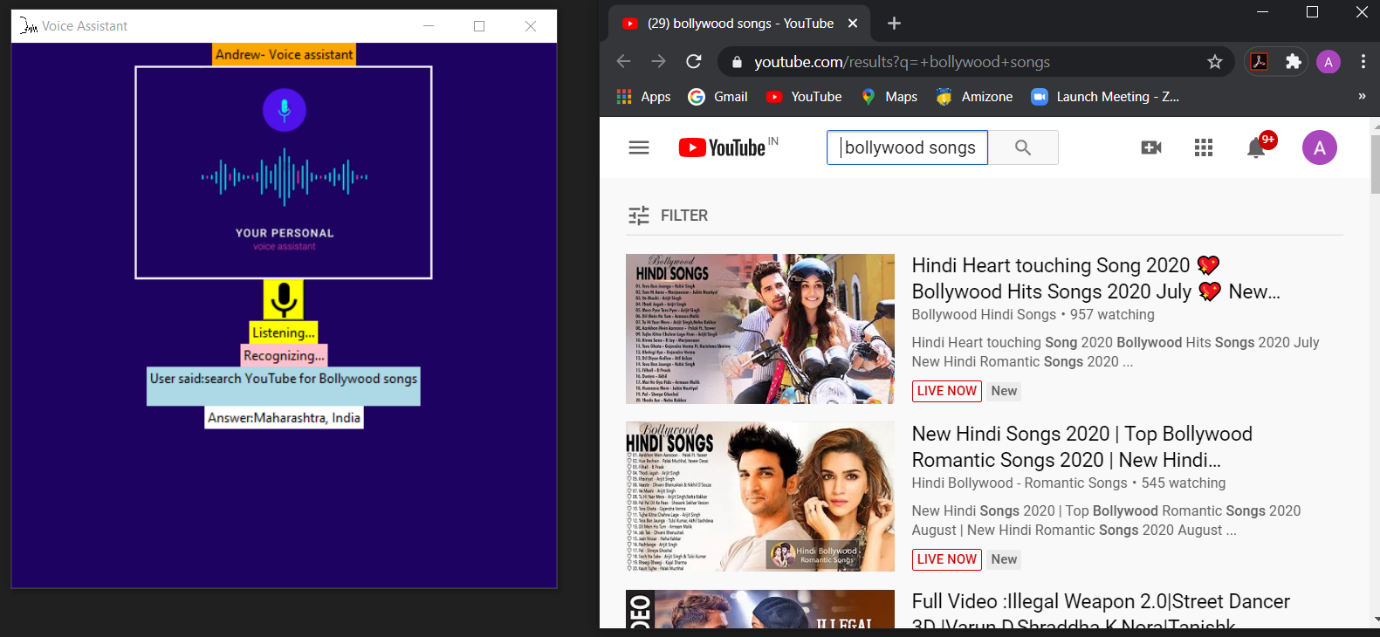
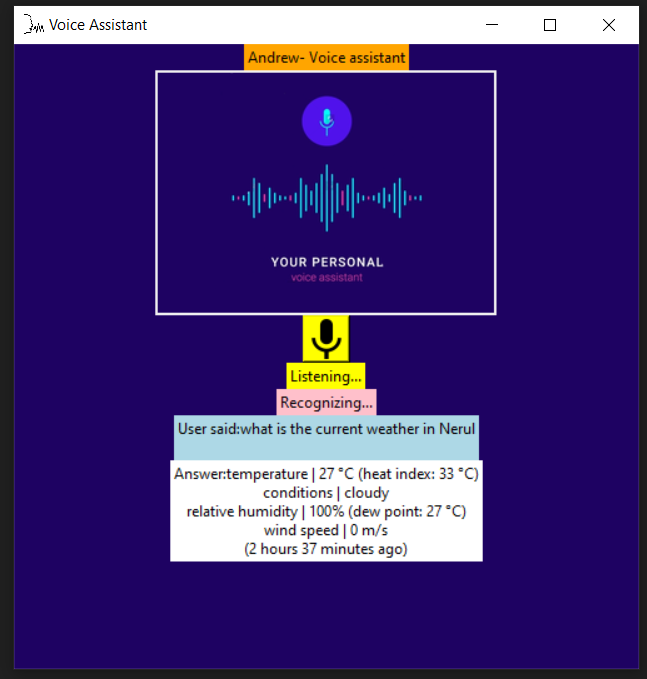












VALIDATION -

So, by determining the above outputs, it is shown that the voice assistant Andrew, can perform all the functions like :- find location, convert currencies of different countries, tell current time, can tell you the latest news, can calculate basic mathematical calculations, can answer your any general knowledge question, mail exchange, event handler, music player service, can tell you jokes, checking weather, Google searching engine, Wikipedia searching engine, direct access and search in popular e- commerce sites like amazon and flipkart, can also do direct search on entertainment platforms like Netflix, Amazon Prime Video and YouTube and initiate the result, can open almost all of the Microsoft office applications like Microsoft PowerPoint, Excel or Word and there are many other tasks which the voice assistant can do just by the voice command of the user.

Although the voice assistant can perform many tasks/functions, we have tried to cover as many screenshots of the output as possible.

**RESULT & DISCUSSION**

The AI voice assistant, Andrew is Designed to help Native and especially for Blind persons which works on their voice commands. Andrew has various functionalities find location, convert currencies of different countries, tell current time, can tell you the latest news, can calculate basic mathematical calculations, can answer your any general knowledge question, mail exchange, event handler, music player service, can tell you jokes, checking weather, Google searching engine, Wikipedia searching engine, direct access and search in popular e- commerce sites like amazon and flipkart, can also do direct search on entertainment platforms like Netflix, Amazon Prime Video and YouTube and initiate the result, can open almost all of the Microsoft office applications like Microsoft PowerPoint, Excel or Word and there are many other tasks which the voice assistant can do just by the voice command of the user.

Hence, Andrew is language barrier independent which actively responds to user’s voice commands faster than the Online Voice Search applications.

Positive Effects with Voice Command-

Voice command can simplify living standards for people living with disabilities, such as reduced hearing, seeing or movability. Furthermore, voice command will help to create a more streamline living for the average person, making chores and purchases for examples much easier. The existence and availability of different safety systems on the market gives the resident the opportunity to protect and keep track of what is going on, increasing the control and safety of the person. Studies show that both voice command and music have the ability to affect human behaviour.

Negative effects with Voice Command-

Constantly talking to a voice assistant will generate vast big data possibilities, providing companies with information that might be used wrongfully by the company- contributing to an uncertainty in the privacy policies. Employing algorithms for user profiling and personalization can have a negative effect on the users, making them feel at unease and suspect or expect security threats. The possibility of numerous jobs at different locations becoming substituted by voice command assistant increases as many simple tasks, such as executing trades, easily could be handled with machine learning software’s as voice assistants and AI. The user relying on the voice command system serviceability too heavily entails a big risk seeing that the voice control is not developed to assist help in potential danger situations.

**FUTURE SCOPE**

Artificial intelligence has truly transformed the way voice assistants are used in our daily lives, and we are only beginning to understand how they will be integrated into all of our activities in the years to come.

I see a time when not only will the algorithms understand what is said, but the way it is said.

This AI Voice Assistant application can be modified in the future to determine the mood of the speaker, whether they are in distress, or how strong or weakly they may believe in the statement. It could be used to adjust the response or, in the case of security situation, whether a response is warranted. It may also lead to a means of diagnosing mental state. Ultimately, performance concepts like sarcasm would be identifiable and play a role in the response.

So basically, this AI Voice Assistant application after these modifications, thus can have a very bright future scope which can be beyond limits.

**REFERENCES & BIBLIOGRAPHY**

[1]<https://www.activestate.com/blog/how-to-build-a-digital-virtual-assistant-in-python/>

[2] <https://docs.python.org/3/library/tkinter.html>

[3] <https://www.geeksforgeeks.org/python-randint-function/>

[4] <https://pyttsx3.readthedocs.io/en/latest/engine.html>

[5] <https://pypi.org/project/SpeechRecognition/>

[6] <https://pypi.org/project/pyjokes/>

[7] “Speech recognition with flat direct models,” IEEE Journal of Selected Topics in Signal Processing, 2010.

THE END!!