AI BASED DESKTOP VOICE ASSISTANT

By

Ankan Biswas

(B.Tech 2nd Year, Enrollment No.: 12022002001040)

Ayush Saha

(B.Tech 2nd Year, Enrollment No.: 12022002001059)

Sankalp Singh

(B.Tech 2nd Year, Enrollment No.: 12022002001138)

Under the Supervision of

Santanu Basak

Dept. of Computer Science & Engineering

University of Engineering & Management, Jaipur

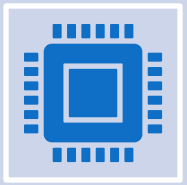
Outlines

- Introduction
- Objectives
- Proposed Model
- Experimental Setup
- Result Analysis
- Conclusion & Future Scope
- Reference
- Acknowledgement

Introduction

Desktop voice assistants are software used to do certain tasks for the user, improve user experience, increase user productivity, etc. In short it is a software made for its user's ease.

Typically, an Intelligent Personal Assistants will answer queries and perform actions via voice commands using a natural language user interface.



Virtual or Desktop assistants are in trend today because everyone wants that their work should be done efficiently with minimum stress and should have excellent quality. So, the virtual assistant is gaining an exponential popularity.

Objectives

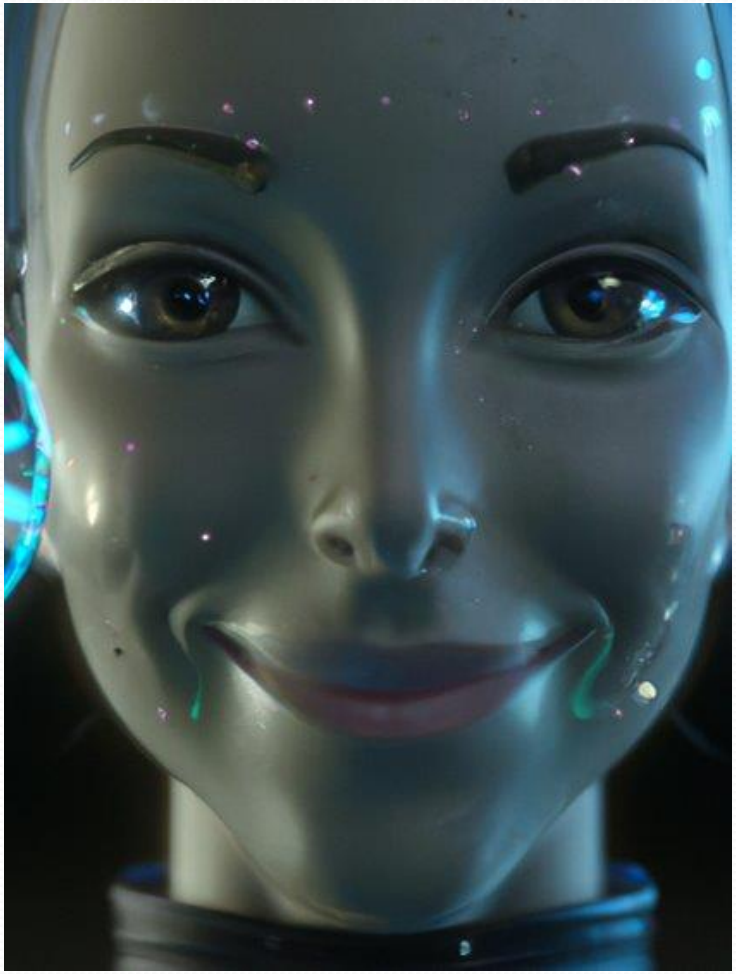
This basic idea behind creating this very Virtual assistant is to create standalone program that helps to the user to perform various task like

- Knowing what is the date and time
- Getting some specific information from Wikipedia
- Searching about something in the web browser
- Opening a website directly from the web browser and even playing some video from the YouTube website.
- Opening a preinstalled application and closing an opened application
- Terminating its own program.

This desktop voice assistant can do all of this by interacting with the user just using the user's voice and can also reply in voice.

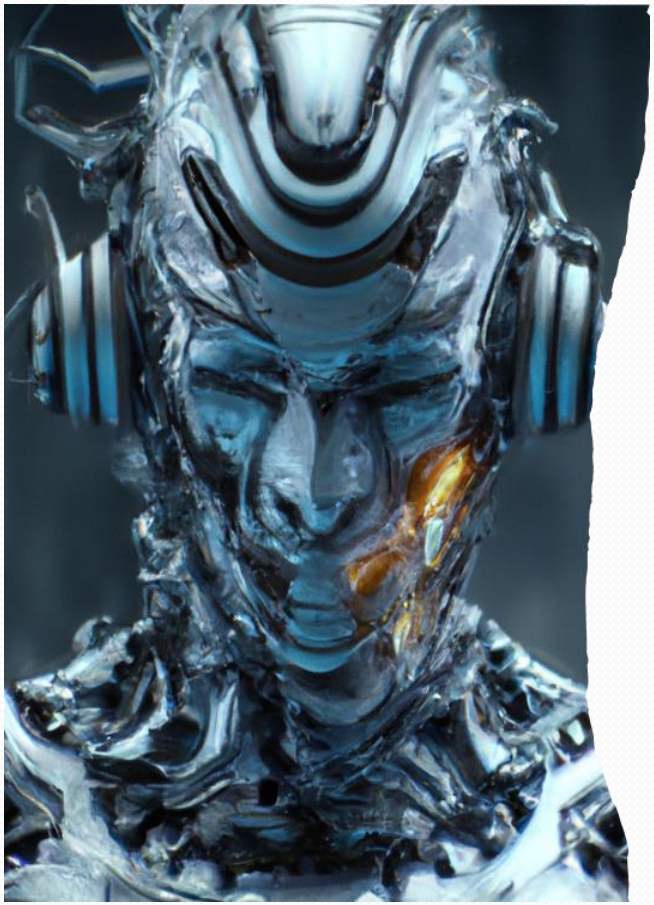
Proposed Model

- The proposed Desktop voice assistant is a program interacts with the user and performs tasks assigned to it though it is not as advanced as other voice assistant like Google Assistant, Alexa, Siri, it can do normal interaction and normal functions like accessing the Wikipedia, searching on the internet, opening website, and so on.
- This was achievable with the help of various module which



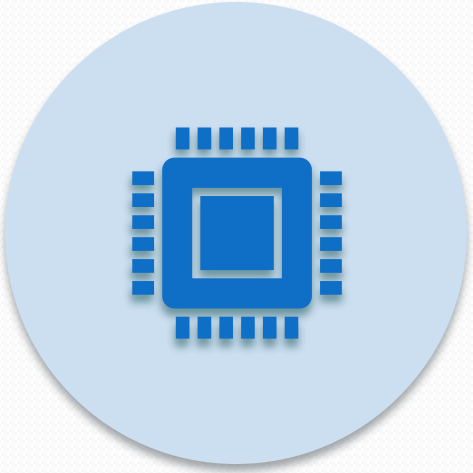
Text-to-Speech Conversion With pyttsx3

- pyttsx3 is a Python library that provides a cross-platform interface for text-to-speech conversion. It supports multiple TTS engines, including SAPI5 on Windows .
- Developers can use pyttsx3 to convert text into speech and customize the voice, speed, and volume of the output. This library can be used in conjunction with SpeechRecognition to create voice assistants that can both recognize and generate speech.



Speech Recognition with Python

ONE OF THE KEY COMPONENTS OF A VOICE ASSISTANT IS SPEECH RECOGNITION. WITH PYTHON, DEVELOPERS CAN USE THE SPEECHRECOGNITION LIBRARY TO CONVERT AUDIO INPUT INTO TEXT THAT CAN BE PROCESSED BY THE PROGRAM.



THE LIBRARY SUPPORTS MULTIPLE SPEECH RECOGNITION ENGINES, INCLUDING GOOGLE SPEECH RECOGNITION AND MICROSOFT BING VOICE RECOGNITION. DEVELOPERS CAN ALSO TRAIN CUSTOM MODELS FOR IMPROVED ACCURACY AND PERFORMANCE.

Wikipedia Package

- *We need to use Wikipedia library so that we can get information from Wikipedia on any topic, or we can also ask for solution to our query or simply we can perform Wikipedia search for any topic using this library. This Library in python needs Internet connection for fetching results and it will provide results to user in text as well as voice format.*

webbrowser module:

- The **webbrowser** module is a convenient web browser controller in the Python programming language. This module offers a high-level interface that enables showing the documents based on the web. Under most circumstances, we can call the **open()** function from the **webbrowser** module to perform the right thing.

Result Example

- Upon running the program, it first greets us and start to listen to our command

```
PS F:\ankan biswas> python -u "f:\ankan
I am Veronica, your Little assistant.
Please tell me how can i help you.
Listening...
```

- If the keyword “how are you” is found in the query or the input speech it responds with

```
Listening...
Recognizing...
User Said : hey how are you

I am fine, Thank you
How are you, Sir
Listening...
Recognizing...
User Said : well it's good
```

- If the keyword “Wikipedia” is found in the query or the input speech it searches the query using the Wikipedia module and responds with

```
Listening...
Recognizing...
User Said : search from Wikipedia what is the Alzheimer's disease

A neurodegenerative disease is caused by the progressive loss of structure or function of
neurons, in the process known as neurodegeneration. Such neuronal damage may ultimately in
volve cell death. Neurodegenerative diseases include amyotrophic lateral sclerosis, multip
le sclerosis, Parkinson's disease, Alzheimer's disease, Huntington's disease, multiple sys
tem atrophy, and prion diseases.
```

Result Analysis

If the keywords “open” and “website” is found in the query or the input speech it opens the web browser and opens the specific website.

If the keyword “exit” or “bye” is found in the query or in the input speech then the program terminates itself.

```
Listening...
Recognizing...
User Said : exit the program

Thanks for giving me your time
PS F:\ankan biswas>
```

Applications of Voice Assistants

- Voice assistants have a wide range of applications, from controlling smart home devices to providing customer service in call centers. They can also be used to improve accessibility for individuals with disabilities, allowing them to interact with technology using natural language.
- In the future, voice assistants may become even more integrated into our daily lives, with the potential to revolutionize the way we interact with technology and each other.

Challenges in Developing Voice assistants

DEVELOPING A VOICE ASSISTANT CAN BE CHALLENGING DUE TO THE COMPLEXITY OF THE TECHNOLOGY INVOLVED. SPEECH RECOGNITION ALGORITHMS MUST BE ABLE TO ACCURATELY RECOGNIZE A WIDE RANGE OF ACCENTS AND DIALECTS, WHILE NATURAL LANGUAGE PROCESSING ALGORITHMS MUST BE ABLE TO UNDERSTAND THE NUANCES OF HUMAN LANGUAGE.

IN ADDITION, PRIVACY AND SECURITY CONCERNS MUST BE TAKEN INTO ACCOUNT WHEN DEVELOPING VOICE ASSISTANTS. DEVELOPERS MUST ENSURE THAT USER DATA IS PROTECTED AND THAT THE VOICE ASSISTANT CANNOT BE USED TO ACCESS SENSITIVE INFORMATION.

Limitation

- While voice assistants have many benefits, they also face several challenges and limitations. One of the biggest challenges is ensuring that the assistant can accurately understand and interpret user requests, which requires sophisticated speech recognition and NLP capabilities.
- Another challenge is maintaining user privacy and security, as voice assistants often collect sensitive information about users. Additionally, there are limitations to what voice assistants can do, as they are reliant on APIs and services provided by third-party providers.

Future Scope

- **Enhanced Natural Language Understanding**
  - Improved context comprehension for more conversational interactions.
- **Personalization & Context Awareness**
  - Tailored responses based on user history & individual preferences.
- **Multi-modal Interactions**
  - Integration with AR/VR for immersive experiences.
- **Advanced Task Automation**
  - Handling complex tasks and decision-making processes.
- **IoT Integration**
  - Seamless control of smart homes & IoT devices.
- **Accessibility & Inclusivity**
  - Catering to diverse user needs and disabilities.

Conclusions

- In this report we have discussed about AI based desktop voice assistant using python. It performs basic tasks like telling date, time, simple interacting with the user, search Wikipedia, searching on the web browser,opening a specific website or opening or closing a specific app.
- The functionality of the program is limited, it can only run and perform its task when it is online, and lacks other functionality also. A fact that is to be mentioned is this voice assistant has no GUI and based on terminal.
- Though this is only a miniature version of a full-fledged Voice assistant it has a lot of scope in future and can be integrate with advanced AI like chat GPT 4, Google Bard, Bing Chat or it can itself become something like that in place of integrating with those AIs.

References

- https://pypi.org/project/pyttsx3/
- https://pypi.org/project/SpeechRecognition/
- https://docs.python.org/3/library/datetime.html
- https://pypi.org/project/wikipedia/

Acknowledgement

The endless thanks goes to Lord Almighty for all the blessings he has showered onto me, which has enabled me to write this last note in my research work. During the period of my research, as in the rest of my life, I have been blessed by Almighty with some extraordinary people who have spun a web of support around me. Words can never be enough in expressing how grateful I am to those incredible people in my life who made this thesis possible. I would like an attempt to thank them for making my time during my research in the Institute a period I will treasure. I am deeply indebted to my research supervisor, Professor Santanu Basak me such an interesting thesis topic. Each meeting with him added in valuable aspects to the implementation and broadened my perspective. He has guided me with his invaluable suggestions, lightened up the way in my darkest times and encouraged me a lot in the academic life.

UEM
श्रद्धावान लभते ज्ञानम्
UNIVERSITY OF ENGINEERING & MANAGEMENT
Good Education, Good Jobs