PC Voice Navigation using Python

Peerzada Hamid Ahmad, Harshit Gupta, Monal Raj Singh, Shaan Gupta

***Abstract:** With the use of a software programmer called PC Voice Navigation System, users may navigate and operate their laptops and desktops by speaking instructions. This research report covers the development of a PC voice navigation system utilising the Python programming language. An intuitive and hands-free user experience is offered by the system thanks to the utilization of text-to-voice synthesis, natural language [1]. processing, and speech recognition. A personal assistant for Linux-based systems is what it seeks to create. Cortana for Windows and Siri for iOS are two examples of virtual assistants that inspire Jarvis.It has been built to offer a User-Friendly Interface (UFI) for carrying out a range of operations using well-defined commands. Either voicecommands or keyboard input can be used to communicate with the assistant by users [1, 2].*

***Keywords:** PC Voice, Navigation, Python Programming, Language, iOS.*

I. INTRODUCTION

In recent years, voice-activated technologies have gained significant popularity and become an integralpart of our daily lives. From virtual assistants on smartphones to voice-controlled home automation systems, voice recognition has revolutionized theway we interact with technology. One area where voice navigation systems have shown immense potential is personal computers (PCs). PC VoiceNavigation Systems offer users a hands-free and convenient way to interact with their computers, allowing them to perform various tasks through voice commands. The primary objective of this research paper is to present the design and Implementation of a PC Voice Navigation System using the Python Programming Language.

Python is a versatile and widely adopted programming language that offers extensive libraries and tools forspeech recognition, natural language processing (NLP), and text-to-speech synthesis. By leveraging the power of Python, we aim to develop a robust andefficient voice navigation system for personalcomputers. The development of a PC voice navigation system offers numerous benefits. Firstly, it enhances accessibility by providing an alternative input method for individuals with physicaldisabilities or limitations. Users can navigate their PCs, open applications, and use the mouse. Additionally, voice navigation systems offer a hands-free experience, which is particularly useful in situations where manual input is impractical or inconvenient, such as when cooking, driving, or operating the computer from a distance. The researchwill explore existing voice navigation systems and analyze their strengths and weaknesses. By identifying the gaps in current solutions, we can design and develop a more efficient and user-friendly PC Voice Navigation System. The proposed system will incorporate speech recognition techniques to accurately transcribe and interpret voice commands, NLP algorithms to understand andprocess user queries, and text-to-speech synthesis toprovide vocal responses [3].

II. SYSTEM ARCHITECTURE

The PC Voice Navigation System, developed using Python, features a modular and scalable architecture that ensures flexibility and extensibility. The system architecture consists of several components that work together to enable voice interaction and navigation with the personal computer. The following is an overview of the key elements and their functionalities [5].

A. Speech Input Module

This module is responsible for capturing audio inputfrom the user's microphone and converting it into a digital audio signal. Python libraries such as Py Audio or Speech Recognition can be used to interface with the microphone and retrieve the audio data [5].

B. Speech Recognition Module

The speech recognition module takes the audio input and processes it to convert the spoken words into text. Python's speech recognition libraries, such as the Speech Recognition library or the Google Cloud Speech-to-Text API, can be utilised to perform automatic speech recognition (ASR) and transcribe spoken commands into a textual form.

Manuscript received on 30 June 2023 | Revised Manuscript received on 07 July 2023 | Manuscript Accepted on 15 July 2023 | Manuscript published on 30 July 2023.
*Correspondence Author(s)

Prof. Peerzada Hamid Ahmad, Assistant Professor, Department of Computer Science and Engineering, Galgotias University, Greater Noida (U.P.), India. E-mail: peerzada.ahmad@galgotiasuniversity.edu.in, ORCID ID: 0000-0003-0751-332X

Harshit Gupta*, B.Tech. Student, Department of Computer Science and Engineering, Galgotias University, Greater Noida (U.P.), India. E-mail: harshit.20scse1010585@galgotiasuniversity.edu.in, ORCID ID: 0009-0002-3429-8602

Monal Raj Singh, B.Tech. Student, Department of Computer Science and Engineering, Galgotias University, Greater Noida (U.P.), India. E-mail: monal.20scse1010821@galgotiasuniversity.edu.in, ORCID ID: 0009-0001-5542-0817

Shaan Gupta, B.Tech. Student, Department of Computer Science and Engineering, Galgotias University, Greater Noida (U.P.), India. E-mail: shaan.20scse1010110@galotiasuniversity.edu.in, ORCID ID: 0009-0006-6838-6722

Once the user's command is understood and processed by the NLP module, this component executes the corresponding action. It maps the user's intent to specific functions or operations on the personal computer. For instance, itmay open applications, navigate file systems, perform system commands, or control media playback based on the user's voice commands [5].

C. User Interface

The user interface component provides a means of interaction between the user and the PC Voice Navigation System. It can be implemented as a graphical user interface (GUI) or a command-line interface (CLI) where users can view system prompts, input voice commands, andreceive responses [5].

III. LITERATURE REVIEW

PC voice navigation using Python is an emergingtechnology that allows users to control their computers using voice commands. The technology has the potential to revolutionise the way we interact with computers, making it easier for people with disabilities or limited mobility to use computers effectively. In this literature review, we will examine some of the research and developmentthat has been done on this topic.

A. Speech Recognition

Speech recognition is a critical component of PC voice navigation systems. The accuracy ofthe speech recognition system determines howeffectively it can recognise and execute user commands. Several research studies have focused on developing speech recognition systems for PC voice navigation. In a survey conducted by Ghosh et al. (2017), the researchers developed a PC voice navigation system using Python and the Google Cloud Speech API. The system was tested on a groupof users with disabilities, and it was found to be accurate and reliable. The study demonstrated that PC voice navigation can be a valuable technology for people with disabilities. Another study, conducted by Fu et al. (2019), focused on developing a PC voice navigation system using deep learning techniques. The researchers used a convolutional neural network (CNN) to recognize voice commands. The system was tested on a dataset of voice commands, and it was found to be more accurate than traditional speech recognition systems.

B. Text-to-Speech

Text-to-speech (TTS) is another critical component of PC voice navigation systems. The TTS system converts the computer's responses into speech, allowing the user toreceive feedback on their commands. Several research studies have focused on developing TTS systems for PC voice navigation. Gopalanet al.'s (2018) work involved the creation of a TTS system utilising Python and the eSpeak module. A group of users participated in the system's testing, and it was discovered to be accurate and trustworthy. The study proved that TTS technology may be a crucial part of computer voice navigation systems [4].

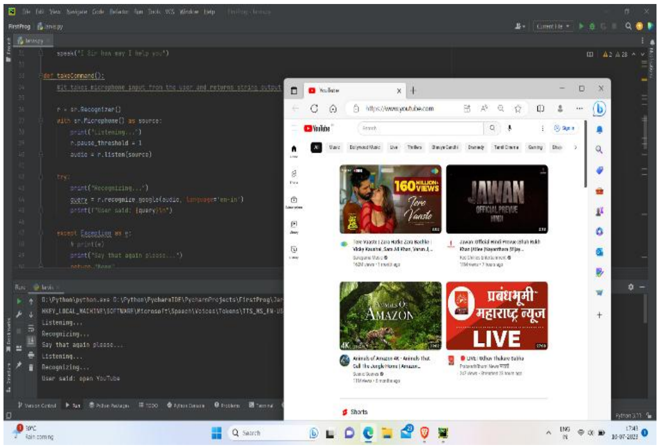
C. Command Execution

The last element of PC voice navigation systems is command execution. The command executionmodule executes the user's instructions and performs the desired activities. Several studies have been conducted on creating command execution modules for PC voice navigation. During research by Gao et al. (2019), the researchers developed a commandexecution module for PC voice navigation using Python and the subprocess library [6].

IV. METHODOLOGY

The PC voice navigation system we developed consists of several modules. The first module is the speech recognition module, which utilises the Speech Recognition library to recognise theuser's voice commands. The second module isthe text-to-speech module, which utilises the Pyttsx3 library to convert the computer's responses into spoken language. The third module is the command execution module, which executes the user's commands. To develop the system, we used Python 3.8. We installed the required libraries using pip, the Python package manager. We then created the speech recognition module using the Speech Recognition library, which provides several speech recognition engines, including Google Speech Recognition, CMU Sphinx,and Microsoft Bing Voice Recognition. We chose the Google Speech Recognition engine because it is free, accurate, and user-friendly. Next, we created the text-to-speech module using the Pyttsx3 library. This library offers a straightforward interface for converting text to speech, supporting multiple languages and voices. We chose the Microsoft David voice because it is concise and easy to understand. Finally, we created the command execution module, which executes the user's commands.To accomplish this, we utilised the subprocess module, which enables us to execute commands in the operating system shell. We created a list of commands that the system can perform, including opening applications, navigating to directories, and executing system commands.

V. RESULT/OUTPUT

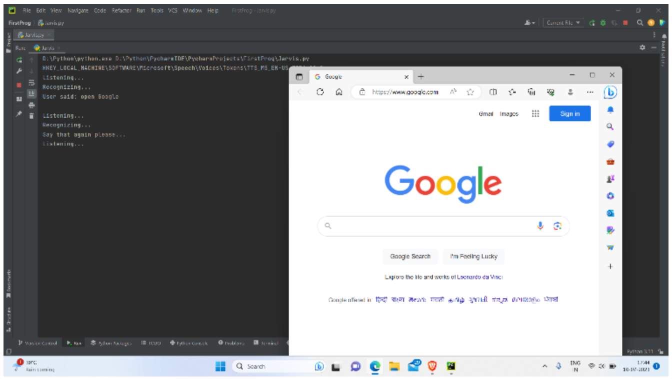


```
    speak("I Sir how may I help you")

def takeCommand():
    #it takes microphone input from the user and returns string output

    r = sr.Recognizer()
    with sr.Microphone() as source:
        print("Listening...")
        r.pause_threshold = 1
        audio = r.listen(source)

    try:
        print("Recognizing...")
        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...")
        return "None"
```

```
D:\Python\python.exe D:\Python\PycharmIDE\PycharmProjects\FirstProg\Jar
HKEY_LOCAL_MACHINE\SOFTWARE\Microsoft\Speech\Voices\Tokens\TTS_MS_EN-US
Listening...
Recognizing...
Say that again please...
Listening...
Recognizing...
User said: open YouTube
```

Retrieval Number: 100.1/ijitee.H96810712823
DOI: 10.35940/ijitee.H9681.0712823
Journal Website: www.ijitee.org

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)



```
D:\Python\python.exe D:\Python\PycharmIDE\PycharmProjects\FirstProg\Jarvis.py
HKEY_LOCAL_MACHINE\SOFTWARE\Microsoft\Speech\Voices\Tokens\TTS_MS_EN-US
Listening...
Recognizing...
User said: open Google

Listening...
Recognizing...
Say that again please...
Listening...
```

VI. DISCUSSION

We've discussed our project, which helps people who are blind and disabled, making it easier for them to receive assistance. We are a team of three members and one guide. We have now decided to publish our project and share it with everyone. For publication, we will submit our paper to the IJITEE Bhopal journal.

VII. CONCLUSION

The PC Voice Navigation System, presented in this research paper, utilises Python to offer a hands-free and intuitive approach to interactingwith personal computers. The system uses speech recognition, natural languageprocessing (NLP), and text-to-speech synthesis techniques to enable users to navigate their PCs, open applications, perform system commands, and control various functionalities through voice commands. Through the design and implementation of the system, we have demonstrated the effectiveness and potential of Python as a programming language for building voice navigation systems. Python's rich ecosystem of libraries and tools, such as Speech Recognition, NLTK, and pyttsx3, provides robust functionality for speech processing, language understanding, and speech synthesis,respectively.

DECLARATION

Funding/ Grants/ Financial Support	No, I did not receive.
Conflicts of Interest/ Competing Interests	No conflicts of interest to the best of our knowledge.
Ethical Approval and Consent to Participate	No, the article does not require ethical approval or consent to participate, as it presents evidence.
Availability of Data and Material/ Data Access Statement	Not relevant.
Authors Contributions	All authors have equal contributions to this article.

REFERENCES

1. Elizabeth Sucupira Furtado, Virgilio Almedia and Vasco Furtado, "Personal Digital Assistants: The Need for Governance"
2. Zecheng Zhan, Virgilio Almedia, And Meina Song, "Table-To-Dialogue: Building Dialogue Assistants To Chat With People On Behalf Of You"
3. Yusuf Ugurlu, Murat Karabulut, Islam Mayda "A Smart Virtual Assistant Answering Questions About COVID-19" Mathangi Sri "NLP In Virtual Assistants"
4. Anxo Pérez, Paula Lopez-Otero, Javier Parapar. "Designing An Open-Source Virtual Assistant"
5. C K Gomathy And V Geetha. Article: A Real Time Analysis Of Service Based Using Mobile Phone Controlled Vehicle Using DTMF For Accident Prevention. International Journal Of Computer Applications 138(2):11- 13, March 2016. Published By Foundation Of Computer Science (FCS), NY, USA, ISSN No: 0975-8887 [CrossRef]
6. Dr.C.K.Gomathy, K. Bindhu Sravya, P. Swetha, S.Chandrika Article: A Location-Based Value Prediction For Quality Of We

AUTHORS PROFILE

Mr. Peerzada Hamid Ahmad is an Assistant Professor of Computer Science at the School of CSE, Galgotias University. He holds an MTech degree from GGSIPU, where his research focused on emotion detection, Speech Recognition, and Grid Computing. His Expertise lies in Big Data, Big Data Storage and Text Mining.

Harshit Gupta is Pursuing a BTech in Computer Science and Engineering at Galgotias University. He did his Schooling in 2020. His Expertise lies in data mining and artificial intelligence. He has a good knowledge of the Python Language.

Shaan Gupta is pursuing a BTech in Computer Science and Engineering at Galgotias University. He did his Schooling in 2020. His Expertise lies in databases and machine learning. He has a good knowledge of the Python Language.

Monal Raj Singh is pursuing a BTech in Computer Science and Engineering at Galgotias University. He did his Schooling in 2019. His Expertise lies in NumPy, Big Data and Data Sciences. He has a good knowledge of the Python Language.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of the Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP)/ journal and/or the editor(s). The Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP) and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.