International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
Article Received: *25 July 2023* ***Revised:*** *12 September 2023* ***Accepted:*** *30 September 2023*

Savvy Virtual Assistant : Transforming the Way We Use Computer

Pallavi Patil
Department of Computer Science and Engneering
Annasaheb Dange College of Engieering and Technology
Ashta, India
pallavipatil2805@gmail.com

Archana Chougule
Department of Computer Science and Engneering
Annasaheb Dange College of Engieering and Technology
Ashta, India
chouguleab@gmail.com

Suhel Sayyad
Department of Computer Science and Engneering
Annasaheb Dange College of Engieering and Technology
Ashta, India
suhelsayyad2006@gmail.com

Ashvini Patil
Department of Computer Science and Engneering
Annasaheb Dange College of Engieering and Technology
Ashta, India
patil.ashvini04@gmail.com

Abstract—Nowadays, technology has a greater impact on daily life and has become more integrated part. Artificial intelligence is playing a bg role in automating daily rouine tasks. Voice assistants developed using AI are helping in integrating daily life with technology. There are a few voice assistants that we are already familiar with, including Siri and Cortana. It can now function as a simple daily schedule reminder, calculator, note writer and search tool in the voice assistant system. This paper demonstrates voice assistance for using personal computer. It details assistance provided using voice input by producing speech output and displaying text on the screen. The major goals of developed voice assistants are to educate users and provide quick, accurate and easy accesss of personal computer.results. The voice assistant will receive input via microphone, translate input speech into computer-understandable language, and provide the user with the necessary solutions and answers to their queries. It supports sending emails without typing a single letter, running Google searches without opening a browser and carrying out a variety of other routine works like launching your preferred IDE and playing music. Each task can be done with just one input command.

Keywords – technology integration, virtual assistance, artificial intelligence, voice assistance

I. INTRODUCTION

Technology has become undividable part of human life. The integration of voice assistants into our daily lives is indeed a significant development in technology.

In the ever-evolving landscape of technology, the integration of voice assistants has sparked a remarkable transformation in the way we interact with computers. With names like Siri and Cortana becoming household terms, these digital companions have seamlessly woven themselves into the fabric of our daily lives. Beyond their initial novelty, these voice-activated systems have evolved to become powerful tools that simplify tasks, enhance accessibility, and redefine our relationship with technology

Voice assistants are designed to receive voice input from users through a microphone and process it into a format that computers can understand. They can then generate speech output to provide answers and solutions to user queries. Voice assistants enhance convenience by allowing users to perform tasks with just a spoken command, reducing the need for typing or manually interacting with devices. A desktop-based virtual assistant is proposed int this paper. It is a software that responds to voice or text commands from the user to carry out a variety of tasks. Because they can streamline daily tasks, save time, and boost productivity, virtual assistants have gained popularity in recent years. A desktop-based virtual assistant has countless potential uses, from automating tedious chores to enabling hands-free access to data and services. With this project, we

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
***Article Received:** 25 July 2023 **Revised:** 12 September 2023 **Accepted:** 30 September 2023*

want to build a virtual assistant that can boost a user's productivity and make their computing experience better overall.

This paper explores the profound impact of voice assistants on the realm of computer usage. We delve into the capabilities that these intelligent systems offer, from their ability to receive voice input and provide speech output to the diverse range of functions they can perform. We will also discuss their role in educating users and delivering quick, accurate results, all while bridging the gap between human communication and machine interaction.

The paper describes building a virtual assistant that will take commands from the user and perform tasks accordingly. The available models of virtual assistants for computers tend to perform only web searches, the proposed virtual assistant enhances these traditional functionalities of virtual assistants while also providing access to desktop applications. This would help the end users to perform tasks through voice commands in a much easier and faster way. The proposed work will provide a desktop-based virtual assistant. This assistant would easily work through multiple desktop applications as well as multiple web applications. The virtual assistant provides a direct response to queries of the users, saves time, is user-friendly, and is easy to access desktop applications.

Moreover, at the heart of this transformation lies the magic of Natural Language Processing (NLP), a technological marvel that enables computers to understand and respond to human language in a more natural, conversational manner. We will journey through the fascinating world of NLP and its implications for voice assistance in computing.

As we embark on this exploration, it becomes evident that the integration of voice assistants into our computer usage opens up new horizons, promising greater convenience and an enriched digital experience. Let us dive deeper into this realm where spoken words shape our technological world.

II. LITERATURE SURVEY

There are a few applications that we have come across during this survey of existing systems. Cortana [1] is a virtual assistant provided by Microsoft. Cortana uses Bing search engine to carry out functions like creating reminders and answering questions for the user. Cortana supports different languages like English, Portuguese, French, German, Italian, Spanish, Chinese, and Japanese. Cortana has some of the lacunas as follows: Cortana is unable to close applications that are opened. Cortana is not able to Shut down and Restart PC. Cortana can't do operations on videos like play, pause, next, full screen, mini player, etc. Cortana can't start typing anywhere on user's demand. It cannot open websites and only performs web scraping.

Mycroft [2] is an open-source artificial intelligence (AI) personal assistant that can be installed on many different gadgets, including Windows laptops and smartphones. Mycroft is made with privacy in mind. It is a strong alternative to proprietary AI assistants like Cortana because it does not gather or keep personal data. One of Mycroft's unique features is its flexibility. The adaptability of Mycroft is one of its distinctive qualities. Because it is open-source, developers can alter it to meet their own requirements. Depending on their preferences, users can also pick and choose which functions to enable or disable. Mycroft can comprehend voice commands and reply to them thanks to its natural language processing capabilities. As a result, it is capable of a wide range of duties, including setting reminders, playing music, delivering weather updates, and conducting web searches. Mycroft has a smaller user base compared to Cortana.

Braina [3] is a virtual assistant powered by artificial intelligence that aims to increase productivity. You can control your computer and compose text using voice commands rather than a typing machine thanks to a variety of innovative features. The use of Braina is advantageous for everyone who spends a lot of time on a computer and tries to avoid tedious typing activities. Some drawbacks of Braina are as follows: An Internet connection is required for voice tasks. It is not suitable for users with serious speech impediments or hearing problems. Punctuation for non-English languages can be erratic.

Mohd Talib Ansari et al. [4] have proposed voice-activated personal assistant built on Python. Here the proposed system (virtual assistant) operates online and carries out routine duties like weather updates, music streaming, Wikipedia searches, opening desktop programs, and so forth. The functionality of the system can only be used online.

V. Geetha et al. [5] have described a Voice-Enabled Personal Assistant for Pc using Python. The paper describes the use of a speech-enabled personal assistant, which can help people save a lot of time by allowing them to carry out tasks with voice commands. Depending on the needs, it was built by utilizing technologies like Speech-to-Text and Text-to-Speech, as well as integrated with additional capabilities.

A goal-oriented technique for agent development called Goal Net is proposed by Jonathan Leung et al. [6]. It provides the foundation for the virtual assistant model that is suggested. The Multi-Agent Development Environment (MADE) 2019 was used to develop the Goal Nets in this work.

Iosif Itkin et al [7]. provide details on a workflow for data-driven services with an universal data analysis procedure. The developed system is solely intended for usage by software testers and is not intended for other users.

N. Sripriya et al. [8] have proposed a speech-based virtual travel assistant for visually impaired people. A speech-based travel bot that can serve as a virtual tour guide is the proposed system. By interacting with the user and supplying pertinent information, the bot will act as a tour guide by proposing locations and offering information about the location, such as opening hours, rating, and address, to help the user learn more about the area. The suggested system can also be quite helpful in circumstances where a person would feel uneasy, like while driving. Passenger information system with a virtual assistant is proposed by Pranay Madasi et al. [9]. Instead of having the user search in the app for the needed information, this virtual assistant can provide voice assistance. Instead of using predetermined schedules or timetables, the same technologies are used on public transportation buses. Real-time bus tracking makes use of a variety of technologies, including GPS, to track the location in real-time and estimate when buses will arrive at particular stations along the route. The people can make efficient use of their time and arrive at the bus stop shortly before the bus comes when this information is disseminated via wired or wireless media. In addition, they can plan their route and choose an alternative mode of transportation if the bus is running late. Mirwan et al. [10] propose use of deep learning technique for developing virtual assistant. The generative model virtual assistant developed using LSTM is a smart virtual assistant who can have conversations on a broad topic. This methodology does not use prepared responses, in contrast to the retrieval-based

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
***Article Received:** 25 July 2023 **Revised:** 12 September 2023 **Accepted:** 30 September 2023*

model, which uses repositories to construct an answer based on a decision tree. They start fresh with every response. Smart virtual assistants based on machine learning and python libraries are proposed by some more authors [12], [13], [14], [15], [16]. They have described the tasks supported by virtual assistants and methodologies used for the same.

III. METHODOLOGY

This section describes methodology used for developing a virtual assistant for computer usage. Speech recognition is the most important part of virtual assistant system. Authors use Hidden Markov Model (HMM) algorithm for speech processing.

A Hidden Markov Model (HMM) is a statistical model used in various fields, including speech processing. In the context of speech processing, HMMs are particularly useful for tasks like speech recognition and synthesis. In speech processing, the underlying idea behind using HMMs is that speech can be modeled as a sequence of states, where each state represents a particular phoneme, word, or sound in the spoken language. These states are hidden, meaning that we can't directly observe them. Instead, we observe a sequence of acoustic features, such as the spectrum of the speech signal. Each state in HMM has associated probabilities for transitioning to other states. These transition probabilities capture the dynamics of speech. For example, some sounds are more likely to follow others in the natural flow of speech. At each state, we generate observations (acoustic features) with associated probabilities. These observations represent the likelihood of a particular acoustic feature occurring given the current state. Given an audio input, HMMs are used to find the most likely sequence of states (phonemes or words) that could have produced the observed acoustic features. This is done using algorithms like the Viterbi algorithm.

It consists of feature extraction and feature matching. The purpose of the feature extraction module is to transform the speech waveform into a representation for additional processing and analysis. This extracted data is referred to as a feature vector.

Signal-processing front-end module performs the conversion of the voice signal to the feature vector. When a voice sample from an unknown speaker is scored against an acoustic model using feature matching, the model with the highest score wins, and the result is taken as a recognized word. HMM can be implemented using various models like Gaussian, Multinomial, GMM, etc. The proposed system makes use of Google Voice transcription for converting speech into text. Google Voice transcription uses Gaussian Mixture Model (GMM) acoustic models for speech recognition. Gaussian Mixture Model can be implemented in python using 'hmm.GMMHMM model' of 'hmmlearn' library.

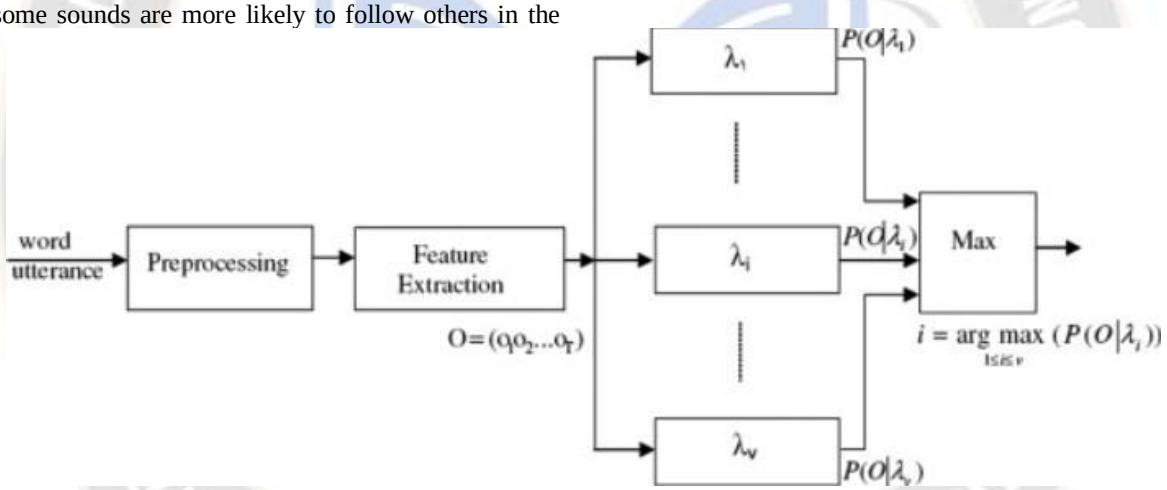

Figure 1. Architecture of the proposed system

The system has been developed as follows: The user provides input to the system in the form of voice commands. The pyttsx3 python package will be used to take voice input from the microphone of the device. This voice input is provided as input to the next module. The voice input is then converted into computer-understandable language through NLP techniques in order to perform further operations. The speech recognition python package will be used for recognizing the input provided by the input module. The output of this module is given as input to the next module. The processed input is converted into a query which is nothing but the task at hand or the operation to be performed. Here keywords from the input are taken into consideration and respective operations are taken to be performed. The output of this module is given as input to the next module. The query provided by the query generator is taken into consideration and the query is processed i.e., the required task or operation is executed. Python packages like web browser, requests, and pywhatkit will be used for processing the queries. The output of this module is given as input to the next module. The processed query or task is taken and the desired output of the voice command is generated and displayed as well. The pyttsx3 python package will be used to provide voice output via speaker or any other audio output device if required.

The system would be in waiting for state once the assistant is opened. The system will take in the input from the user and if the user gives a query, then the query would be executed. If the input is a command, then the command would be executed. The result in both alternative situations would be given as output. After the output is provided, the system will again have two alternative workflows. If the user chooses to exit the system, then the Sleep command would be provided and the system would stop. Else after completion of the task, the system would again go back to the waiting state, and the whole process repeats.

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
Article Received: *25 July 2023* ***Revised:*** *12 September 2023* ***Accepted:*** *30 September 2023*

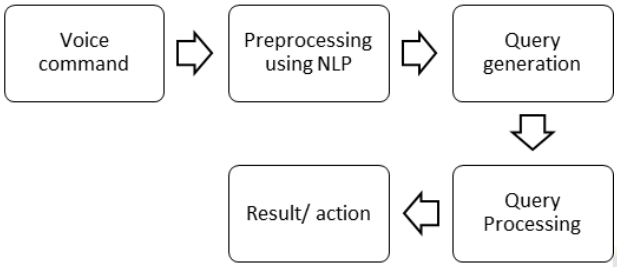

Figure 2. Flow of the proposed system

The system is developed using python programming language. Important python libraries used for developing system are pyttsx3 and Speech Recognition. Python text-to-speech conversion can be done via Pyttsx3. It provides a simple way to convert written text into spoken words. The library uses the native text-to-speech engines of the operating system, which means that it can produce high quality speech output in various languages and voices. Speech_recognition is a library for performing speech recognition, with the support of several engines and APIs, online and offline.

IV. RESULTS

The results obtained from the implementation of the proposed methodologies are as described to get more insight into the final results. Some of the project's goals, such as gathering voice input from the user, evaluating that voice input, and executing some queries based on input, are successfully met. The assistant carries out operations like starting and stopping notepad and command prompt, obtaining the user's IP address, shutting down and restarting the computer, sending messages via WhatsApp to the static user, obtaining the day's news, getting the local time, and starting the camera.

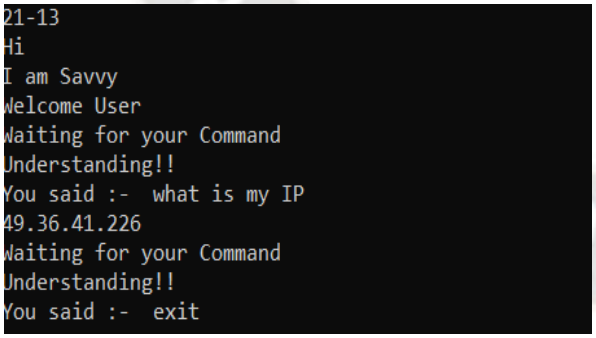

Figure 3. Command processing by Virtual assistant

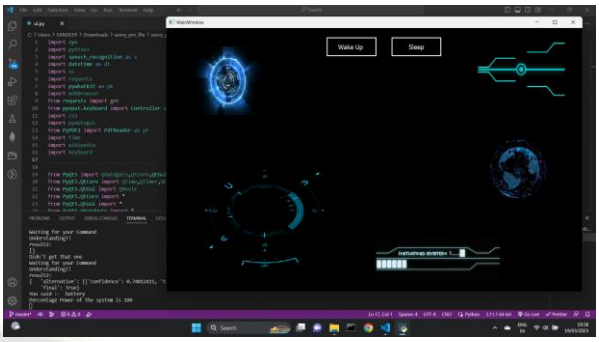

Figure 4. Voice processing by virtual assistant

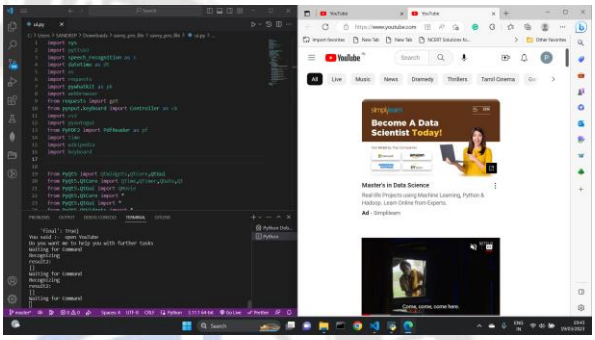

Figure 5. Internet access using virtual assistant

Word Error Rate (WER) is a metric used for evaluating the performance of a Speech Recognition system. WER compares an actual statement with a predicted statement.

$$WER = (S+D+I)/N$$

where S is the number of substitutions, D is the number of deletions, I is the number of insertions and N is the number of reference words.

$$W\ Accuracy = 1- WER = 1 - (S+D+I)/N = (N-S-D-I)/N$$

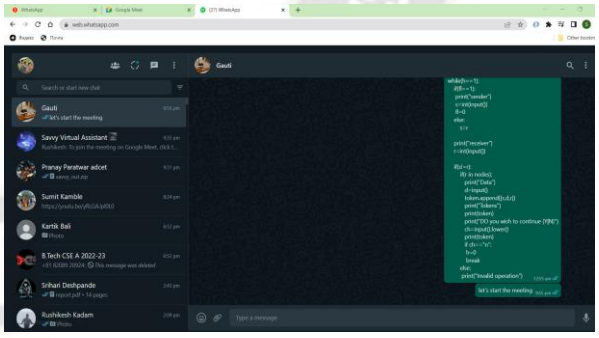

Figure 6. WhatsApp access using virtual assistant

For finding out accuracy of voice to text conversion using different input devise, system has been tested with system microphone and headphones. Graph shows the comparison between the accuracy of speech-to-text conversion while using headphones and while using the system microphone for Harvard sentences which are used for standardized testing of cellular, Voice over IP, and other telephone networks.

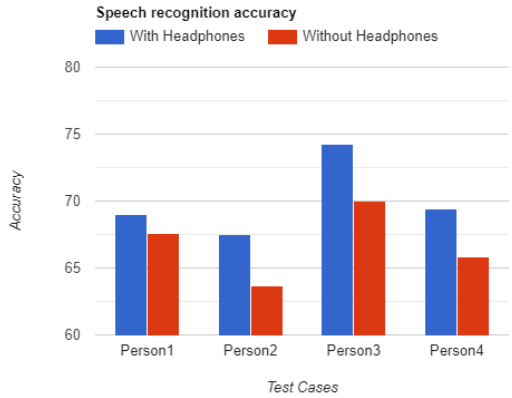

Figure 7. Accuracy of speech to text conversion

Limitations of the Savvy virtual assistant are as follows: It is less secure and does not provide any encryption methods for securing data. Voice recognition accuracy varies based on the microphone quality of different desktop devices. Most of the tasks require internet connectivity. Unlike google chatbot, it cannot stop working in between any tasks.

V. CONCLUSION

The proposed system provides voice-based assistance to use personal computer. The system answers almost all questions asked by the user. A simple user interface is given for starting and closing the virtual assistant. Savvy virtual assistant performs tasks like opening different applications and websites, and typing dictated tasks when required. It also provides answers to user questions like what is, who is, how to etc. Some of the project's goals, such as receiving voice input from the user, processing that voice input, and executing queries based on input, are successfully met. The assistant carries out operations like opening and closing various apps, shutting down and restarting the PC, sending messages via WhatsApp to the static user or sending mail to anyone, obtaining today's news, getting the local time, starting the camera, doing google searches, opening websites, and general answers, etc.

So, it is evident that desktop-based virtual assistants will continue to develop in order to satisfy our always-changing needs. These virtual assistants will become more and more crucial in assisting us in navigating our increasingly digital future. So that we can maximize its advantages while limiting any potential drawbacks, it is crucial that we embrace new technology and use it responsibly. Future Scope Functionalities like person detection and voice recognition can be added. Authenticated user access can also be added. The system can also be modified to access complex desktop apps. More security can be provided by using encryption methods. Better voice recognition could be done using various noise cancellation algorithms.

REFERENCES

[1] Heena Reyaz Bhat, Tanveer Ahmad Lone, Zubair M. Paul, "Cortana-Intelligent Personal Digital Assistant: A Review", International Journal of Advanced Research in Computer Science, pp. 66-57, 2017
[2] Matthew Scholefield, "Mycroft Documentation" , Release 0.1.0, ttps://buildmedia.readthedocs.org/media/pdf/mycroft-core/release-v19.2.7/mycroft-core.pdf, pp. 1-30, 2019
[3] "Braina", https://www.brainasoft.com/braina/download.html [Accessed on 03/11/2023]
[4] Mohd Talib Ansari, Ansar Ahmed, Aadil Jiwani, Prajaktee Rane "Desktop Virtual Assistant" International Research Journal of Engineering and Technology (IRJET), Volume: 08 Issue: 05, May 2021
[5] V. Geetha, C.K.Gomathy, Kottamasu Manasa Sri Vardhan, Nukala Pavan Kumar, "The Voice-Enabled Personal Assistant for Pc using Python" in International Journal of Engineering and Advanced Technology (IJEAT) ISSN: 2249-8958, Volume-10 Issue - 4, April 2021.
[6] Jonathan Leung Jonathan Leung, Zhiqi Shen, Chunyan Miao, "Goal-Oriented Modelling for Virtual Assistants" 2019 IEEE International Conference on Agents (ICA)
[7] Iosif Itkin, Andrey Novikov, Rostislav Yavorskiy, "Development of Intelligent Virtual Assistant for Software Testing Team" 2019 IEEE 19th International Conference on Software Quality, Reliability and Security Companion (QRS-C)
[8] N. Sripriya, S.Poornima, S.Mohanavalli, R.Pooja Bhaiya, V. Nikita, "Speech-Based Virtual Travel Assistant For Visually Impaired" 2020 4th International Conference on Computer, Communication and Signal Processing (ICCCSP)
[9] Pranay Madasi, Karthik Ponnam, Prashanth Mandalapu, "Passenger information system with a virtual assistant" 2017 International Conference on circuits Power and Computing Technologies [ICCPCT]
[10] Mirwan, Aryo Nugroho, Ferial Hendarta, Rumaisah Hidayatillah, Firdaus Hassan, Kristovel Printo Nana "VIRTUAL ASSISTANT USING LSTM NETWORKS IN INDONESIAN" 2018 International Seminar on Research of information technology and intelligent systems (ISRITI)
[11] https://pyautogui.readthedocs.io/en/latest/keyboard.html [accessed on 02//02/2023]
[12] Pavani Kollamudi, Bhargavi Koduru, Bharathi Poranki, Madhu Priya Golla, "Smart virtual assistant", Materials Today: Proceedings, Vol. 80, pp. 3575–3578 , 2023
[13] Smita Srivastava, Devesh Katiyar, Gaurav Goel, "Desktop Virtual Assistant", International Journal for Research in Applied Science & Engineering Technology (IJRASET), Vol. 10, Issue III Mar 2022
[14] Zeeshan Raza, Aftab Amin Sheikh, Ankush Shahu, Abhijeet Thakur, "Virtual Desktop Assistant Using Python", International Research Journal of Modernization in Engineering Technology and Sci., Vol. 04, Issue 04, April 2022
[15] Vijaya Balpande, Vedanti Lute, Neha Pawar, Aayush Jain, Saniya Sadaf, "Virtual Desktop Assistant Using Machine Learning", IJARIIE, Vol-8 Issue-2 2022
[16] Shubham Thorbole, Anuradha Pandit, Gayatri Raut, Tejas Sirsat, "Ai-Based Desktop Voice Assistant", Journal of imerging technologies and innovative research, Volume 10, May 2023