

Voice based Intelligent Virtual Assistant for Windows

Sheetal Phatangare

Assistant Professor
Department of Computer Engineering
Vishwakarma Institute of Technology
Pune, Maharashtra

**Shardul Kedar, Sakshi Shewale,
Asif Shikalkar, Avani Shilimkar**

Department of Computer Engineering
Vishwakarma Institute of Technology
Pune, Maharashtra

ABSTRACT

Nowadays there are various improvements in the field of Automation. Developing technology which will reduce human efforts without changing the accuracy of the end output is the main goal of the developers in this field. Keeping the same goal in mind the proposed system. 'Voice Based Intelligent Virtual Assistant for Windows' focuses on reducing human efforts by taking commands in voice format given by the user. The definition of a Voice Assistant is obtained from progresses in artificial intelligence and speech recognition as well as natural language processing. The developed system which we name as 'JK' is built using python which is one of the most popular programming languages in the field of Artificial Intelligence and Machine learning. By including multiple libraries like subprocess, pytsx3, json, random, operator, speech recognition as sr, datetime, Wikipedia etc, system is capable of doing a wide range of tasks by taking user's voice as in input command. The system is capable of taking pictures, opening different applications, playing music, opening files located at different locations, taking notes, changing wallpapers, having random chats with the user and performing many more functions.

KEYWORDS: *Python; Virtual assistant; Speech recognition; Desktop; Personal assistant; Voice assistant; Artificial intelligence*

INTRODUCTION

In the past events, humans depended on other humans for help or services. The globalization of technology ensured that today, people don't need to ask anyone else for aid, they may rely on a gadget that is far more effective, dependable and also can take care of their daily requirements.

Speaking of the modern-day situation, absolutely each person desires to keep itself away from doing boring work. To this end, we've got created a project called "Virtual Voice based assistant for Windows" that lets us to control our duties surely via way of means of speaking a command. We are actually applying this concept in many tasks like Google's assistant, Siri and Cortana, etc.

The activity of a voice is described in 3 stages: 1. text to speech; 2. Intension to speech; 3. Intension to action. Virtual Personal Assistant has nearly grown to be a primary necessity in all digital gadgets if you want to

execute the desired problems easily. More than simply being a bot, VPA could make lifestyles easier for the consumer in numerous ways. AI-based language assistants are useful in many areas IT help desk, home automation, HR tasks, etc. Language-based searches will continue, to be the future for the next generation of people.

Given the usefulness of PVA, its use density is increasing rapidly. For example, 21% of the US population owns at least one smart speaker and 81% of adults own a smartphone. Therefore, it is very likely that the user is always within range of at least one of his PVA. Virtual desktop assistant that we have developed is using Python as it offers various libraries that we can use it for making a virtual assistant. Sapi5 and Espeak of Linux may prove helpful for voicing our machines.

LITERATURE REVIEW

The proposed system conducts a personal voice assistant security and privacy survey. The main aim

was to focus on developing measures which will make the voice assistant more secured to use. Since all the commands are received in the form of voice, privacy issues may rise as voice recordings may reveal sensitive user information.

The main focus was developing an authentication system which will differentiate between the user and an attacker.

The following research proposed an Artificial Intelligence based virtual assistant where the main focus of their system was to develop a simple to use Voice User Interface(VUI). The system they proposed was capable of opening different applications, searching on google. It can even extract the current location of user and send it to user's family in case of any emergency.

A Virtual Assistant using Artificial Intelligence is proposed in this paper. The main goal of the system is to execute more commands as compared to other Voice assistant in the market. But one of the limitations of the project was the time taken to execute the commands was comparatively higher than other assistants. Complex algorithms with higher time complexity are used which is one of the main reasons which leads to more time consumption while executing different commands.

A systematic literature review on Virtual Assistants for Learning has been put forth in this research paper. The main focus of the paper was to study and analyze the role of virtual assistant(VA) in the field of higher education in order to understand how VA can help students in time management and other aspects of education. It also covers points regarding different techniques which can help to develop a VA.

The proposed has deeply explained the concept of Desktop Assistant. The paper focuses on the main challenges faced by the desktop assistant which need to be addressed in order to build an efficient desktop voice assistant. Future scope of the system is also discussed under which the authors have mentioned improvements required in the existing system. One of them being interfacing the desktop assistant with IOT technology which will revolutionize the Automation industry.

This essay examines how artificial intelligence is becoming more prevalent in daily life, particularly in

the form of voice assistants that can understand natural language.

Using open-source software modules, the design and implementation of a digital assistant are discussed, with an emphasis on flexibility and the simplicity of adding new features. The assistant relies on verbal input and tries to relieve the user of pointless manual labor. Ultimately, the article demonstrates the potential for voice assistants to promote natural human-machine interaction.

Voice control technologies and virtual assistants are quickly evolving, with increasingly sophisticated cognitive computing capabilities enabling for multilevel requests and complex tasks. Virtual assistants, being programs based on cloud, need internet-connected hardware or software to run, and the data they produce can be used for machine learning and artificial intelligence. The proposed system can ease interactions with other programs and modules while supporting organization, and it has the potential to transform how we engage with our gadgets and carry out daily chores. This study provides as an example for advanced applications in the automation area.

The study of human and script communication is the subject of natural language processing (NLP), a relatively new field of study in computer science and artificial intelligence. The following research works in the areas of Text-to-speech and automatic voice recognition to improve voice recognition technologies. The study involves computational and numerical modelling of many elements of language.

By employing multi-modal communication channels, such as speech, graphics, video, gestures, and more, the proposed system seeks to improve interaction between people and computers. The system offers a natural dialogue experience by integrating technologies like gesture and picture detection, speech recognition, and knowledge libraries. This technology is applicable in many different industries, such as education, healthcare, robotics, and home automation, among others. The new VPAs model will provide human-human communication a cogent framework, making it an important tool in the quickly expanding field of interactive conversational systems.

In order to address the shortcomings of current approaches, the following study suggests a powerful voice recognition system that makes use of machine learning. The goal is to develop a Virtual Personal Assistant (VPA) that is more useful and effective for everyday use. Although the system is very effective, jobs might take longer to finish than with other VPAs, and future improvements might be difficult due to the intricacy of the algorithms. VPAs are a rapidly expanding field whose main objective of AI is to increase accessibility of human-computer interactions. In order to lower error rates, the suggested system takes a novel approach by combining both visual and audio information.

The proposed report looks at how new technologies might be used to build a smart Virtual Personal Assistant (VPA) that focuses on user-based information. It looks at examples of intelligent software that uses natural language processing and various types of help today. The research makes the case that virtual personal assistants may become a reality soon thanks to new technologies. It offers proof that a fundamental VPA with natural language processing algorithms and the capacity to operate without requiring human input or programming is presently feasible based on experiments and user testing.

The voice-controlled personal assistant described in this study features several functional modules, including voice control, character recognition, and virtual assistance. The assistant has been properly constructed and evaluated and utilizes AI and data analytics to give a personalized user experience. The proposed Raspberry Pi-based assistant offers ease and simplicity for impaired users. The paper emphasizes the potential of voice assistants to improve natural human-machine interaction.

This study proposes an application built for people with communication difficulties, such as autistic children, using a PDA with 120 symbols based on daily life. Users can use the PDA's camera or voice recorder to create their own icons using the application's two modes, Supportive and Self-use. According to the paper, the application might be helpful for people who have a variety of communication challenges and might

promote a society in which people support one another regardless of disabilities.

The construction of a robotic system that the operator may command through speech and gesture inputs is discussed in the paper. The system includes a gripper arm for object manipulation and uses Google text to speech API and Grassfire algorithm for controlling its basic locomotion. The goal is to build a working prototype of an automated personal assistant that can be used in domestic and industrial settings to increase productivity while minimizing human effort. Speech processing, gesture control, and image processing are the three main areas of the robot's operation.

Please embed tables and figures in appropriate areas within the document and center them horizontally. Tables and figures should not exceed the given page margins. Provide captions (maximum length: 6 to 8 words) for each table or figure. Centre the caption above the table and below the figure. Please reference the table or figure in the text (see Table 1). Please do not use vertical lines in tables. For figures, GIF and JPEG (JPG) are the preferred formats.

METHODOLOGY

Python-based voice-based virtual assistants use a variety of Python-built libraries. The main libraries used speech recognition, pyttsx3, Wikipedia, datetime, time, web browser, random, tkinter, wolframalpha, and so on. We chose the libraries based on the requirements.

These libraries function in two ways: the first is to provide data required for the assistant, such as the time and date, and the second is to perform the required tasks in different functions used by the assistant. We created various functions to perform various tasks. These functions take the query as input, which is taken in the form of speech from the user.

The voice-based virtual assistant works by accepting commands from the user using the speech-recognition library in Python. After accepting and recognizing the user's commands, they are converted into a query using pyttsx3, and this query is passed into the various functions we prepared for the assistant's operation. The basic working principle is that we take a spoken command, convert it to text, and then pass that text as input to our program.

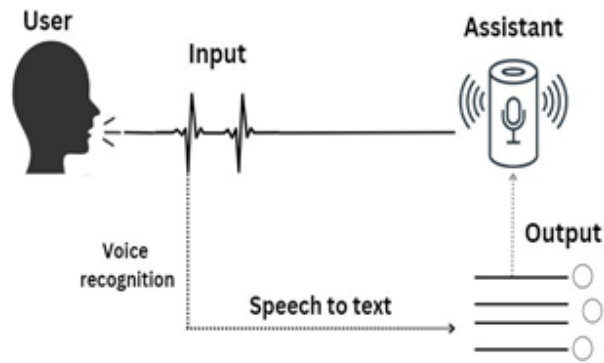


Fig. 1 : Flow of the system

The flow of the system we used in our virtual assistant is depicted (see Fig. 1). The AI looks for keywords within the query and then executes the function. The functions designed to perform various tasks also make use of Python libraries, such as Wikipedia, which is used to provide the user with information about anything directly from Wikipedia in the form of speech without having to navigate to the Wikipedia page through any browser. If we give the assistant permissions over the device, the assistant can also perform functions such as shutting down the system or sending it into sleep mode. The various functions used include playing music, opening different applications (such as PowerPoint, word, excel, etc), taking notes, screenshots, and starting the camera. For opening a particular application or playing a music individual functions are created where we need to pass the folder path of these applications as the parameter of the function. For ex for playing music, we assigned the folder path where the music is located to a variable `music_dir`. Then using the `os.listdir` function we pass `music_dir` as a parameter which returns the songs of the folder and using `os.startfile()` the system plays the song. The function `os.startfile()` available in the `os` library in python is the important function used to open all applications once the path of the application is known and passed as one of the parameters of the function `os.startfile()`.

We also included the function for converting text files to audio using the assistant. The function `speak()` is used to convert a text message into audio. It is used if we want our virtual assistant to react in a particular way in a certain situation. For example, while writing a greeting

function we want our VA to say some sentences like "How are you? "Etc so in this scenario the `speak()` function comes into picture.

The system also allows users to write a note and view the written note. For achieving this we open the file using `open()` function in write and read mode respectively. One of the parameters `open()` function needs is the name of the txt file that needs to be created for writing the note and the mode of the file is the other parameter needed.

The main technologies used in our research are speech recognition and natural language processing. Various algorithms such as PLP features, the Viterbi search, deep neural networks, discriminating training, and the WFST framework, are used by this kind of technology.

A subfield of artificial intelligence, computer science, and human language is known as "natural language processing," or NLP. The development of technology has made it possible for machines to decipher, study, manipulate, and understand human languages. Topic segmentation, giving automatic summary, named entity recognition (NER), audio recognition, and translation are some of the tasks it helps organize knowledge for.

The following table gives us an overview of the features and the drawbacks of similar projects and cited materials.

Table 1 : Comparison of existing Voice Assistant Systems.

Title and Reference	Features	Limitations
AI Based Voice Assistant	Capable of translating multiple languages as specified by user.	Limited command-keyword list, leading to smaller range of executable commands.
Desktop Virtual Assistant	Able to perform a wide range of simple, direct actions like web search.	Unable to perform system related actions like shutting down or restarting.

Study of Voice Controlled Personal Assistant Device	A server for updating data to Firebase cloud. Also, IoT implementation for allowing connection of smart devices.	Complex data flow structure results in more time consumption for each command and hence less efficiency.
Viv - The Personal Voice Assistant	Assistant settings can be changed at runtime	Provides only basic functions to facilitate ease of access for the disabled.

To summarize the features which will overcome the limitations pointed out from the cited material, this system will provide the following functions:

1. Data fetching and decoding structure is simple and hence time required for execution of each individual command is reduced.
2. Includes a wide range of commands and key-word command list which increases the versatility of the VA.
3. Easily perform system related actions which includes shut down, restart, sleep mode, etc.
4. Perform all the basic functions of an assistant in a fast and efficient manner.

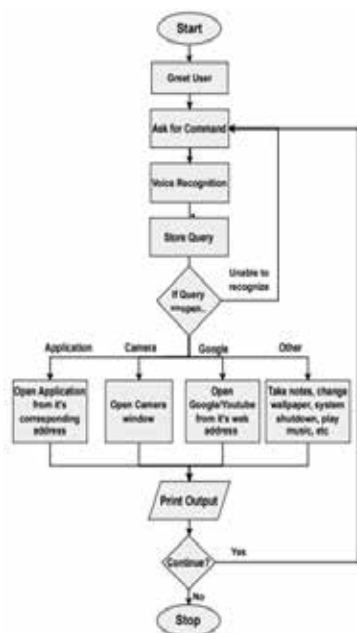


Fig. 2 : Flow of the system

RESULTS

Following are the outputs of various plausible commands that are executable by the assistant:

Table 2 : Accuracy and efficiency of the proposed system

Task	Time taken (in sec)		Average time taken T	Accuracy (Out of 5 attempts)
	T1	T2		
Open Power Point	8	5	6.5	4
Having chat	4.73	6.45	4.54	5
Take a note	4.52	4.56	4.54	5
Play music	4.35	4.65	4.50	5
Show a note	4.53	4.81	4.67	5
Open YouTube	4.25	4.69	4.47	4
Open Google	5.13	15.65	10.39	4
Take a picture	8.94	7.49	8.21	5

T1 = 1st reading of Time required to execute a command

T2 = 2nd reading of Time required to execute a command

T = Average of readings T1 and T2.

The necessary Python programming language packages were installed, and PyCharm's Integrated Development Environment (IDE) was used to write the code. The outputs we have received from our AI-based voice assistant are shown below. The Python code we developed runs in Python 2.7 and Python 3.x. Due to how business has changed over the past few decades, virtual assistants have become more and more common. Many businesses are preferring to invest in flexible work arrangements and schedules rather than spending money on enormous office buildings and overworked personnel. As a result, team organization and project management are now more modern and agile.

DISCUSSION

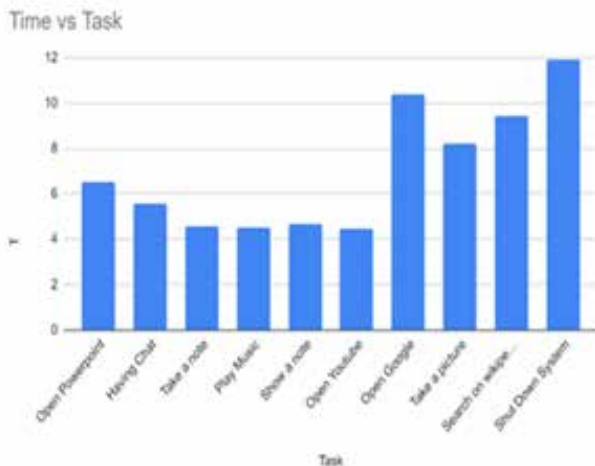


Fig. 3 : Efficiency of the system

(See Fig.3) the Time vs Task i.e., it represents the time required to perform a given task, from recognizing the command to executing it. On observation, we can infer that direct system related tasks, like shutting down the system as well as tasks related to searching on internet are a bit longer as it takes some time for processing the sequence of execution.

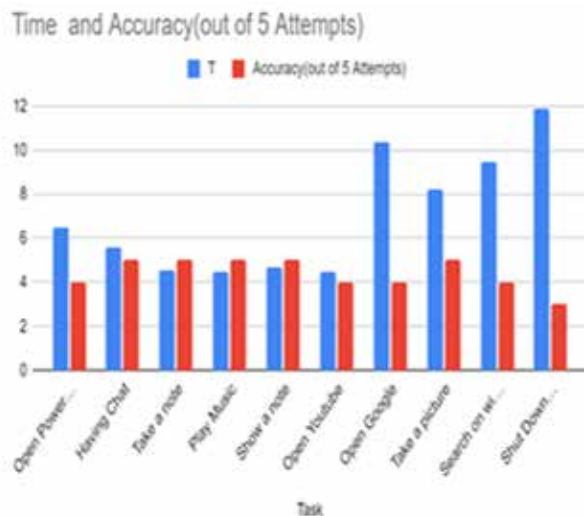


Fig. 3 : Accuracy of the system

(See Fig.4.) It represents the average time required to perform and complete a task and the accuracy with which the task is performed. The readings show that

the accuracy for all the commands is somewhat similar, which depends upon the recognizing capacity of the VA and the clarity of the user's speech.

CONCLUSION

PVA is now ubiquitous and is dramatically changing the way users interact with laptop systems. Customer is increasingly relying on PVA as the primary, or perhaps only, interface to laptop systems and intelligent environments. As a result, the safety of these devices has become the focus of public attention and research efforts. Similarly, privacy is an issue for most users. In our project we have used many other functions compared to other virtual assistants. Voice assistants are helpful in many areas with the education, daily use, consumer electronics, etc. and are becoming more and more advanced in our daily lives.

Many companies are working to improve their language assistants' interactions and upgrading other features, and a lot of young people have started making use of voice assistants in their daily lives from many sources, the results show positive feedback. A lot of advancement is shown in language assistant from the past 2 years. Voice assistants can also be an advantage for the people who can't read or write. Thanks to the AI based voice assistant for paving the way for the users.

REFERENCES

1. Cheng, P., & Roedig, U. (2022). Personal voice assistant security and privacy—A survey Proceedings of the IEEE, 110(4), 476507. <https://doi.org/10.1109/jproc.2022.3153167>.
2. Subash S, Prajwal Srivastava. (2020). AI Based Voice Assistant. Fourth World Conference on Smart Trends in Systems, Security and Sustainability, 593-596.
3. A. Sudhakar Reddy M, Vyshnavi, C. Raju Kumar, and Saumya. (2020). Virtual assistant using artificial intelligence. JETIR, 1116-1119.
4. Regina Gubareva and Rui Pedro Lopes. (2020). Virtual Assistants for Learning: A Systematic Literature Review. CSEDU, 1, 97-103.
5. Smita Srivastava, Dr. Devesh Katiyar, Mr. Gaurav Goel. (2022) Desktop virtual assistant. International Journal for Research in Applied Science & Engineering Technology, 10(3):962-964.

6. Deepak Shende. (2019). AI based voice assistant using python. International Journal of Emerging Technologies and Innovative Research, 6(2):506-509.
7. Chaitanya Kulkarni, Abhay Dekate. (2016). Study of voice controlled personal assistant device. International Journal of Emerging Trends & Technology in Computer Science, 42(1):42-46.
8. Abhijeet Shetty, Aadit Shinde, Yash Sutar. (2022). Viv - the personal voice assistant. International Journal for Research in Applied Science & Engineering Technology, 10(4):6778-6782.
9. Santosh K Behera. (2020). Natural Language Processing for text and speech processing: a review paper. International Journal of Advanced Research in Engineering and Technology, 11(11):1947-1952.
10. Veton Këpuska, Gamal Bohouta. (2018). Next-generation of virtual personal assistants (microsoft cortana, apple siri, amazon alexa and google home). IEEE 8th Annual Computing and Communication Workshop and Conference, 99-103.
11. R. Augustian Isaac. (2018). Virtual personal assistant. Journal of Network Communications and Emerging Technologies, 8(10).
12. Peter Imire. (2013). Virtual personal assistant. Proceedings of 10th Conference of the Italian Chapter of AIS.
13. Vashistha, P., Singh, J. P., Jain, P., & Kumar, J. (2019). Raspberry Pi based voice-operated personal assistant. Proceedings of the Third International Conference on Electronics Communication and Aerospace, 10(4).
14. Torii, I., Ohtani, K., Shirahama, N., Niwa, T., & Ishii, N. (2012). Voice output communication aid application for personal digital assistant for autistic children. 11th International Conference on Computer and Information Science, 329-333.
15. Shaikh, Q., Halankar, R., & Kadlay. (2020). A Voice assisted and gesture-controlled companion robot. Proceedings of the International Conference on Intelligent Computing and Control Systems, 878-883.