Virtual Personal Assistant using RPA, AI and Python

Pranav Gautam, B.Tech. Student, CSE Department, Medi-Caps University, Indore-453331
Vivek Patidar, B.Tech. Student, CSE Department, Medi-Caps University, Indore-453331
Prasanna Kapse, Faculty, CSE Department, Medi-Caps University, Indore-453331

Abstract—**Technological progress is accelerating day by day. Human-computer interaction is becoming more convenient as a result of speech-enabled virtual assistants that provide a wide range of services and can connect to smart environments, reducing physical interaction with the system. However, people face several issues while interacting with these assistants due to their numerous weaknesses, like a lack of understanding complex queries and not maintaining the personalized data of users, etc. As a consequence, the interaction is impaired, and in some cases, ineffective. For this reason, we have developed a virtual assistant using RPA (Robotic Process Automation), AI (Artificial Intelligence), Python, and ML (Machine Learning) that can eliminate these problems and do so with increased accuracy. Hence, this paper proposes to create a virtual assistant that can make human-computer interaction as smooth and efficient as possible.**

Index Terms—**Virtual Assistant, RPA, Artificial Intelligence, Python, Machine Learning.**

I. INTRODUCTION

PEOPLE no longer rely on other humans for assistance or services. Because of the digitalization of the world, humans no longer need to rely on others for assistance; instead, they may rely on considerably more systematic and dependable equipment to meet their daily demands. VPA (Virtual Personal Assistant) has practically become a requirement. However, although it is reasonably efficient, it is not particularly useful and is not used by the user owing to its high mistake rate. Hence, the aim of our project is to create a VPA with a very low error rate.

As our world becomes more digital, virtual assistants powered by advanced AI are bridging the gap between the digital and human worlds. Consumers and corporations may get help with a variety of chores. But what is a virtual assistant, and what can it help you with?

A virtual assistant is a digital assistant that uses voice recognition features and language processing algorithms to recognize the voice commands of the user and perform relevant tasks as requested by the user.

A virtual assistant is capable of doing a variety of tasks. Users can use it to get basic answers and recommendations based on their profile, previous activity, and additional behavior. It may be beneficial in a variety of sectors, including IT helpdesk, home automation, HR-related chores, voice-based search, and so on. Voice-based search is the future of the next generation, with users relying heavily on voice assistants for all of their requirements.

Fig. 1. Virtual Assistant.

Organizations can use this digital assistant in customer care to manage incoming contacts. It can also support a wide range of IT operations. Routine procedures like system upgrades, information management, and even transaction orders may be automated. Virtual assistants are extremely beneficial to the elderly, the blind and physically disabled, children, and others by ensuring that interacting with machines is no longer a barrier for them. Even blind people who can't see the gadget can use their voices to communicate with it. In summary, we've developed an AI-powered virtual assistant that can execute all of these tasks and more without causing any disruptions.

II. PROPOSED WORK

Virtual assistants should be able to model complex task dependencies and use these models to recommend optimized plans for the user. It needs to be tested for finding optimum paths when a task has multiple sub-tasks and each sub-task can have its own sub-tasks. In such a case, there can be multiple solutions to paths, and it should be able to consider user preferences, other active tasks, and priorities in order to recommend a particular plan.

Virtual assistants are the most advanced kind of robotic automation. Our virtual assistant gathers information from a variety of sources and contextualizes it, learning from each encounter. The virtual assistant can process everything that is said or written and utilize it to generate an accurate answer using advanced language processing. Using AI and machine learning, our virtual assistant can handle various tasks and complicated queries. It uses neural networking to acquire insight into one's preferences based on prior selections and data. In this way, when the user interacts with the virtual assistant, it

becomes a unique experience that meets their needs.

This virtual assistant extracts information and complicated data from interactions using advanced artificial intelligence (AI), robotic process automation (RPA), natural language processing, and machine learning. RPA is a smart software application that excels in automating repetitive and rule-based processes using highly organized data. However, RPA can also get confused during the time of a complex query. If data is more unstructured or conventional criteria aren't applicable, cognitive automation can help. The term "cognitive automation" refers to a mix of RPA and data science methodologies that is particularly effective when dealing with textual information. With a speech and text-based user interface, our virtual assistant takes this to the next level. It is capable of extracting information and complicated facts from conversations and comprehending them appropriately.
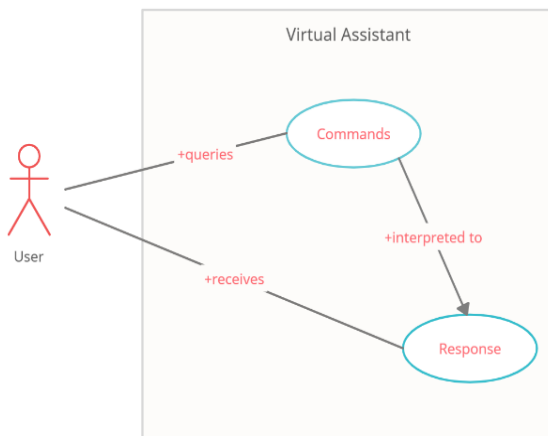

Fig. 2. Use Case Diagram.

The main goal of this research is to improve the accuracy of the voice-to-text software. That is, the software will potentially be able to translate any speech with minor modulations or varied accents into text with the high degree of accuracy and precision required for VPA's day-to-day use.

III. MODELING

1) Speech Recognition
The program converts speech input to text using Google's online speech recognition algorithm. With voice input, users can get texts from specialized corpora arranged on the information center's computer network server, which are briefly kept in the system before being forwarded to Google Cloud for speech recognition. The equivalent text is then received and fed to the central processor.

2) Python Backend
The Python backend receives the voice recognition module's output and determines if the instruction or voice output is an API Request and Context Extraction.

3) RPA
Robotic Process Automation (RPA) is a software technology that allows digital operations like data entering, word processing, and chatting to be automated.

4) Cognitive Automation
It is the application of machine learning that is used to automate processes that would otherwise require physical labor.

5) NLP
Natural language processing is an area of AI that aims to enable computers to interpret text and spoken words in the same manner that humans can.
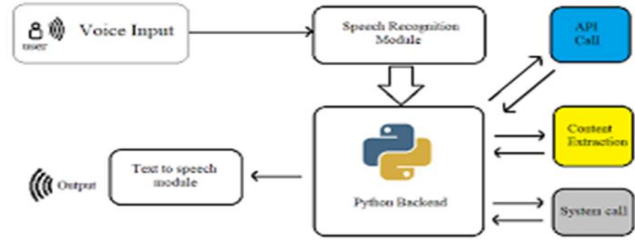

Fig. 3. System Architecture.

6) API Calls
A software interface that allows two programs to interact with one another is known as an API (application programming interface). To put it another way, an API is a messenger that transmits your request to the provider and then returns the response.

7) Content Extraction
Context extraction (CE) is the process of obtaining structured data from machine-readable materials that are unstructured or semi-structured. In most situations, this activity entails employing natural language processing to process human language documents (NLP). Test results for context extraction may be seen in recent operations in multimedia document processing, such as automated annotation and content extraction from images, audio and videos.

8) Neural Networking
The assistant may learn about users and become more customized as a result of neural networking, eventually carrying out activities without the user having to ask.

9) Machine Learning
ML aids the assistant in doing many jobs and answering tough queries. VPA also uses machine learning to acquire insight into a person's preferences based on prior decisions and data.

10) Text-to-Speech Module
The capacity of computers to read text aloud is referred to as text-to-speech (TTS). Written text is converted to a phonemic representation, which is subsequently converted to waveforms that may be generated as sound by a TTS Engine. Third-party publishers offer TTS engines in a variety of languages, dialects, and specialist vocabularies.

IV. FUNCTIONALITY
The proposed system will have the following functionality:
  a) The assistant will be started by the user either manually or by voice activation (the voice activation command can be set by the user, like, hey Petra) and it will listen for the command or query from the user.

b) According to the user's question, the assistant will search for the answer and provide the best solution possible.

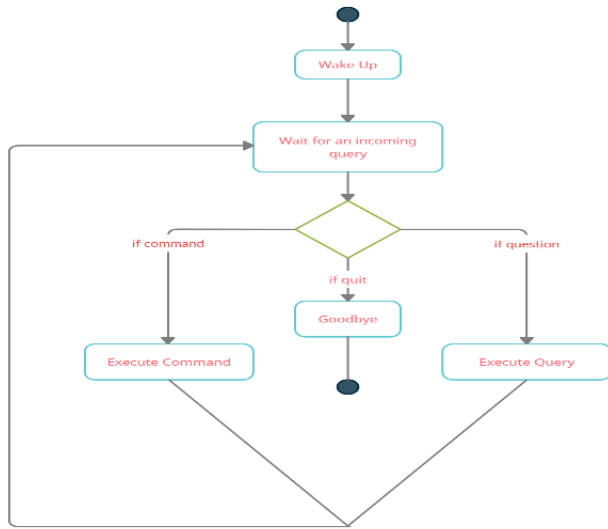

Fig. 4. Activity Diagram.

c) If the assistant is unable to listen, it will again ask the user for some input. This time of asking again can be modified by the user. According to the user's preferences, the assistant can have a male or female voice.

d) If the assistant is unable to extract information from the user's input, it will ask the user to repeat the process until the desired number of times has been reached.

e) Finally, if the assistant doesn't get any input, it will shut itself down. Or if the user is done with it, they can close the assistant just by saying any stop word of their choice (quit, goodbye, exit).

V. RESULT AND ANALYSIS

This section provides a brief description of our result on the basis of the comparison and analysis of our proposed work. We have employed this idea by the means of RPA, AI, Python, and ML.
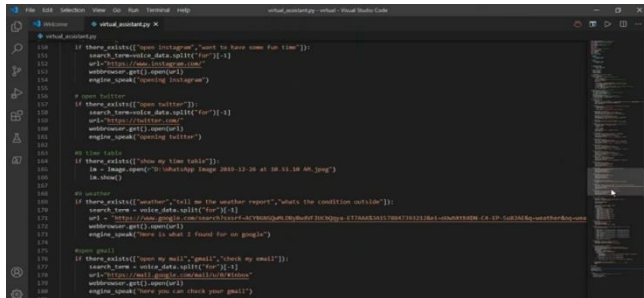

Fig. 5. Windows Code.

This is the Windows code that will be executed on the client side to take the user's voice input.
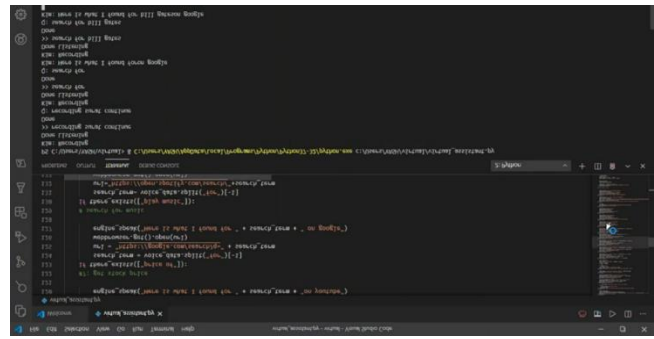

Fig. 6. Assistant taking input through voice (user searching for Bill Gates).

The first output that will be presented when the Windows Code is performed is that the assistant is "recording". The user must then issue the voice command. After this, the assistant prints "done listening".
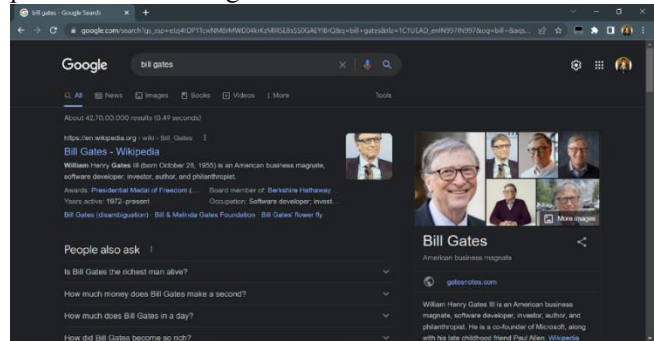

Fig. 7. Output given by the assistant.

After the translation of the command that the user has given, the output will be displayed.

Users may utilize this Virtual Personal Assistant to get an intelligent computer secretariat service. The service is built on the internet's combination of voice recognition, artificial intelligence, and machine learning. The VPA reduces the user's interruptions, improves his time management, and offers a single point of communication for all of his messages, contacts, calendar, and information sources. It is anticipated that this will become a common feature for millions of users.

VI. CONCLUSION AND FUTURE SCOPE

In this project, "Virtual Personal Assistant Using RPA, AI, and Python", we discussed the design and implementation of digital assistance. The project is developed with open-source software components that are supported by the PyCharm community and can quickly accept any modifications. Because this project is modular, it is more adaptable and easier to add new features without disrupting existing system functionality. It not only responds to human instructions but also delivers replies depending on the inquiry or words said by the user, such as opening tasks and operations. It greets the user in a way that makes them feel more at ease and allows them to communicate freely with the voice assistant. The application should also reduce any extra manual labor that is required in the user's day-to-day tasks. The whole system is based on verbal input.

Our virtual assistant can improve efficiency and offer support to both employees and customers of the organization. It allows them to offer more services by taking over more of the routine tasks. Hence, employees can spend more time on other tasks. It not only allows them to offer more and better services, but it also allows the organization to save money. This virtual assistant can increase the safety factor in automobile engineering as it can eliminate the physical interaction of the driver with the machine. They can easily communicate with the VPA through their voice, which can handle almost any query or command.

Virtual assistants are becoming a popular and useful technology, with a variety of advantages, contributing to the automation of tasks and providing support in time-management, accessing information, communication facilitation, etc. The technology is still in its early stages. The Virtual Assistant presented in the paper has few new functionalities compared to other assistants, but an additional and forward-looking feature may be embedded as a future work of this project.

This is only the beginning; as AI and machine learning advance, our virtual assistant will get smarter and provide new possibilities.

VII. REFERENCES

[1] Bassam A. & Raja N. (2010) "Arabic speech recognition using Hidden Markov Model Toolkit (HTK)". 2. 557 - 562. 10.1109/ITSIM.2010.5561391.

[2] ELSHAFEI, M., VIRTUAL PERSONAL ASSISTANT (VPA) FOR MOBILE USERS. MITEL NETWORKS (2000–2002).

[3] B. S. Atal and L. R. Rabiner, "A pattern recognition approach to voiced unvoiced-silence classification with applications to speech recognition," Acoustics, Speech and Signal Processing, IEEE Transactions on, vol. 24, no. 3, pp. 201–212, 1976.

[4] Emad S. Othman. "Voice Controlled Personal Assistant Using Raspberry Pi". International Journal of Scientific and Engineering Research Volume 8, Issue 11, November-2017.

[5] V.Radha and C. Vimala, "A review on speech recognition challenges and approaches," doaj. org, vol. 2, no. 1, pp. 1–7, 2012.

[6] "VIRTUAL ASSISTANT USING ARTIFICIAL INTELLIGENCE", International Journal of Emerging Technologies and Innovative Research (www.jetir.org), ISSN:2349-5162, Vol.7, Issue 3, page no. 1116-1119, March-2020.

[7] George Terzopoulos, Maya Satratzemi, "Voice Assistants and Smart Speakers in Everyday Life and in Education", Informatics in Education, v19 n3 p473-490 2020.

[8] Regina Gubareva, Rui Pedro Lopes, "Virtual Assistants for Learning: A Systematic Literature Review", CSEDU20-RP-93.

[9] Jatu Naazneen, Abdul Gaffar, Abhijit Palse, "Project Report on Virtual Assistant", DTSS COLLEGE OF COMMERCE, 2018-2019.