RESEARCH PAPER

DESKTOP BASED SMART VOICE ASSISTANT USING PYTHON LANGUAGE

Jaya Gupta 2301920140073

Master of Computer Application

GL Bajaj Institute Of Technology And Management (Greater Noida)

[Muskanguptarani@gmail.com](mailto:Muskanguptarani@gmail.com)

***Abstract*-** In today's world, voice assistants, also known as speech-based artificial intelligence (AI), have become indispensable. These applications react to spoken commands, facilitating smooth human-computer interaction. Voice assistants are very popular these days and come in very handy when things get hectic. These programs, like Amazon's Alexa and Google's Assistant, which even young children can use, have become almost ubiquitous, particularly during the global pandemic when smartphone use increased. Such as IoT characteristics can be used to control and manage household equipment for those with mobility issues. Even for those with physical impairments, voice assistants like Alexa are quite helpful because they can perform duties like controlling IoT-enabled devices in addition to providing entertainment.

***Key words:***

***Internet of Things, Python, Speech to Text, Text to Speech, Voice Assistants.***

1. INTRODUCTION

Voice assistants are one of the most well-liked and quickly developing advancements in today's tech-driven society, and they are quickly becoming a necessary component of everyday life. They are really helpful because they can do a lot of things with just a spoken word. Even children who can speak the necessary language, which is typically English, may readily operate them because they only require the user's voice. Voice assistants are particularly helpful for people who have physical limitations. For example, a person who is unable to walk can just tell the assistant to switch off the lights. Voice assistants may help customers keep organized by ordering food, creating reminders, and handling a variety of other activities with ease, in addition to operating IoT-enabled devices.

A voice assistant's capacity to simplify processes and save a substantial amount of time is one of its main benefits. Stated differently, it makes performing multiple tasks smooth.

1. LITERATURE REVIEW

**Aarthi Easwara Moorthy** (2014) et al. With just their voice, customers may operate household equipment thanks to the Voice-Activated Personal Assistant. However, users' interactions with the assistant may be impacted by the presence of outsiders. Additionally, the voice-activated assistant model acts as a basic interpreter, using the desktop assistant to carry out user orders[1].

**Veton Këpuska** (2018) et al. According to the plan for the upcoming generation of virtual personal assistants, voice assistants will improve human-system interaction by combining speech recognition with gesture, image, and video detection. The user will receive the results when the assistant selects the best output device. There are no system calls made throughout this operation [2].

**George Terzopoulos** (2019) et al. recommended smart speakers and voice assistants for daily use and instruction. They asserted that natural language processing, or artificial intelligence, would significantly improve people's daily lives. Its Internet of Things capabilities will be immensely beneficial to blind people, allowing them to carry out everyday tasks [3].

**Subhash S** (2020) et al. the author developed speech-to-text technology using a recorded audio file to create the proposed voice assistant, which is driven by artificial intelligence. The functions are carried out once the audio has been processed. The Google Text-to-Speech (GTTS) engine, which is used in this project, is backed by a strong library. However, it does not offer capabilities like Internet of Things (IoT) integration and system calls [4].

**Benedict D. C** (2020) et al. suggested Artificially intelligent voice assistants will influence consumer choices by examining human behavior and producing more powerful psychological responses. There are Internet of Things characteristics in the assistant. Although there are certain drawbacks to this paper, it may also order anything that the user desires. To keep up with voice conversations, the voice assistant depends on speaker's capabitlity to express available options. Its inability to make system calls is another major drawback [5].

**Nivedita Singh** (2021) et al. idea, which used a Python speech-to-text module in conjunction with APIs and system calls, enabled the creation of a Python voice assistant that enables users to carry out tasks via voice, without the need for a keyboard. Hybrid platforms can also be used with this assistance. The article does have certain shortcomings, though, like insufficient system call support [6].

**Abeed Sayyed** (2021) et al. A presentation was published on the use of Python and IoT features in desktop assistants. The project combined Python with a SQLite database and artificial intelligence (AI) capabilities. While it lacks API call and system call functionality, it includes a records conjunction and the query framework to support its operations [7].

**A.M.Sermakani** (2021) et al. The project suggests developing desktop speech recognition using Python programming. With Python serving as the backend, it integrates Teach-To-Speech, voice activation, voice biometrics, AI technologies, and automatic speech recognition. To improve its overall capabilities, the project also incorporates cutting-edge brain technology[8].

1. PROPOSED WORK

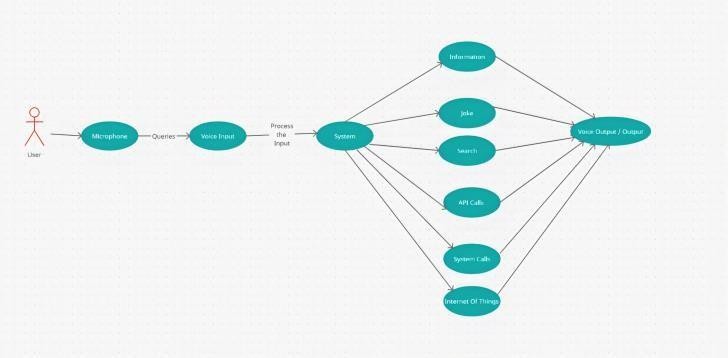
One of the most important issue solvers in the current world is the voice assistant. Every problem's answer can be quickly and simply determined in a matter of seconds using only the user's voice; no more information is needed. It's also helpful for

Figure 1 Activity Diagram

1. **Geetha** (2021) et al. One voice command can be used to turn off, restart, or recite the most recent news with the Voice-Enabled Personal Assistant for PC, a Python project that the author demonstrated. This helper project's backend is written in Python. This project has a well-supported library for converting raw JavaScript Object Notation (JSON) data into text, even if not all APIs can do so. Nevertheless, the processing of request calls is delayed [9].

**Dimitrios Buhalis** (2021) et al. A study on the ways in which voice-activated AI digital assistants in rooms are boosting hotel services and guest experiences was proposed. During the COVID-19 epidemic, when human touch is deemed dangerous, these assistants are especially helpful as they allow guests to utilize hotel amenities via voice commands. Although the lack of human interaction is not viewed as a disadvantage, staff training is necessary and the integration of these systems is complicated. Voice assistants can also be utilized to regulate the temperature and lighting in rooms, giving visitors more convenience and security [10].

house maintenance, like turning on or off anything with a stopwatch. Therefore, in the modern world, a voice assistant is an essential piece of AI software. which calls for "Hello NOVA" (Next-Gen Optimal Voice Assistant) as a wake-up phrase.

Finding a file can be time-consuming and may cause deadlines to be missed. But the search is far faster if you just ask the voice assistant to open the folder. One of the main benefits of employing a voice assistant is the opportunit

In this section our goal is to comprehend the use case diagram design process. The following are some potential situations that the system could run into:

Each user is shown as an actor in a Activity Diagram, and NOVA's voice assistant feature is shown as a Use-Case. The user asks the assistant a question while speaking into a microphone to start the conversation. The user's spoken input is converted into text by the speech-to-text module, which then processes it through a number of modules and takes the necessary action. A voice assistant should have the following essential features*: -*

* 1. Providing Information Upon Request-   
     Any information the user requests must be provided by the voice assistant. Online data reading and searching can take a lot of time, but a voice assistant makes this process easier by swiftly retrieving and displaying the information you need. This illustrates how a voice assistant can help a user save a lot of time.
  2. Providing the user's location with the day's top stories: -

By delivering updates that are customized to the user's preferred location or subject, a voice assistant removes the need for conventional news consumption methods. Watching a news channel might take up a lot of time because it may show unrelated items or news from other places before the user's desired information is covered. This procedure is streamlined by a voice assistant, which only provides the user with the news they desire.

* 1. Cracking Puns to Lighten the Mood: -

Let's discuss what every person has experienced

difficult situations or arguments with close friends and family at some point in their lives. Therefore, a random joke can lighten the mood at least 10% of the time, which could end the argument or cool us down. One quote that is relevant to the ideas discussed in this paragraph is "Laughter is the best medicine."

4) Extracting the desired file or folder:

Everything must be done quickly in our fast-paced environment in order to keep to the schedule, and sometimes people require help to finish chores quickly. However, chores can be completed quickly and without interruption when using a voice assistant. For instance, the voice assistant can instantly retrieve a reference file if the user is working on documentation and subsequently needs it.

* 1. Giving the Client's Location's Weather and Heat Information: -

Why does it matter what the weather is like right now? The explanation is straightforward: it aids users in getting ready for the day. The voice assistant might say something like, "It may rain today, so take an umbrella," or "It will be sunny, so don't forget your sunglasses." This useful feature emphasizes the voice assistant's contribution to users' preparedness and knowledge, making it even another crucial feature.

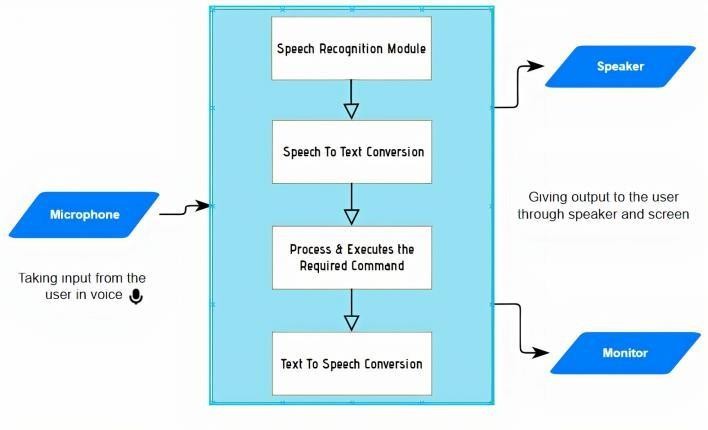
* 1. Looking up the user's query:

People in the 21st century often have questions that need to be answered right now. If they don't, their doubts will only grow larger, and eventually they will have n doubts. To get answers to our questions, we can search the internet, and asking an assistant will save a lot of time. In addition to answering queries, we must conduct extensive online research to stay current with trends. We may accomplish this by simply instructing our assistant to conduct a search on a particular topic or query.

* 1. Internet of Things:

The last and most important feature is  IOT. It turns out to be really beneficial and saves a significant amount of time. Imagine, for instance, that a person who has trouble walking has to turn on a fan, but the switch is too far away. They can ask the voice assistant to turn on the fan in this situation, and it will comply. With the aid of IoT, numerous jobs may be completed, this being just one of them.

Though it provides much more than these features, these are the essential features of a voice assistant. A voice-controlled assistant can be incorporated into many facets of daily life thanks to technological improvements, offering even more efficiency and ease.



**Fig 2. Workflow Model**

The assistant recognizes the wake-up word when the user speaks into a microphone, converts the spoken input into text using Speech to Text (STT), interprets and comprehends the command, and then carries out the

user-specified task. Lastly, the assistant delivers the response in an AI-generated voice using a Text to Speech (TTS) module.

1. *The constraints: -*

Voice based assistants have been around for a while in today's technologically advanced society, and the current systems have few limitations or disadvantages.

.

* 1. The Imperfectness of Speech Recognition– Whenever an end user poses a question, the assistant may misunderstand and provide alternative answers, or the user may need to repeat the question in order for the assistant to comprehend and process it. For instance, it occasionally fails to distinguish between homonyms like "their" vs "there."
  2. Ambient Noise Disruption - The fact that voice assistants require a peaceful atmosphere in order to work efficiently is one of their main drawbacks. They might find it difficult to distinguish between background noise, other conversations, and the user's voice in noisy environments, which could result in misunderstandings or incorrect command processing.
  3. Security Concerns - Questions and details about the device's accounts and services can be retrieved by anybody with access to a voice-activated device. This poses a serious security issue since the gadget might read private information including calendar entries, emails, and other confidential data. Voice assistants are susceptible to a number of dangers, such as man-in-the-middle attacks, which can jeopardize user security and privacy.

1. EXECUTION OF SUGGESTED WORK
2. *The purpose of Python:*

We have selected Python as the language for the helper's development since it supports Object-Oriented Programming and has a number of built-in functions that make the process easier. The assistant's processing of inquiries can be tailored to the needs of the user. In order for the assistant to comprehend and react to user requests, speech recognition entails translating audio into text. Python is perfect for this project because of its adaptability, which enables it to be utilized for a variety of jobs. Due to its increasing popularity, it is now used in sophisticated domains such as natural language processing, artificial intelligence, machine learning (ML), and data science. Numerous libraries are available in Python to satisfy every need of this project.

Although these are the voice assistant's key features, there are many other things we can do with it. The helper has the following features available for use:

-Displaying any video that the user desires to watch

-A random knowledge shared at the beginning of the day can teach the user something new and motivate them to approach their work in a more interesting and educational manner.

One feature that every assistant should have is the ability to engage in gaming so that users may spend the time in a fun way.

* With a voice assistant, we can turn off the system even after we've left the location where it is by just telling the assistant to do so. Users could neglect logging off the machine, which could cause important data to be lost or left unsaved.

1. *How is this Speech Assistant proficient ?*

A basic summary of the functions that must be included in voice assistants is provided below, as described.

API CALLS

Using API keys, we have retrieved news articles from the internet using the News API, a straightforward REST API based on JSON. This enables the assistant to search for the most recent information on particular subjects or display the most recent news from multiple websites. Additionally, well-known APIs like OpenWeatherMap are used to receive weather forecasts. The assistant can also provide detailed weather information using neural machine learning, allowing users to make decisions based on up-to-date information from anywhere in the world. This data can be reliably retrieved and displayed to the user by the system.

SYSTEMCALLS  
  
We have utilized the OS & Web Browser Module to access a number of system functions, including the desktop, calculator, task manager, command prompt, and user folder. Additionally, this module enables the assistant to restart the machine and open the Chrome application as necessary.

DATA RETRIEVAL

Through the use of Selenium's WebDriver module, which offers a variety of web driver implementations, the assistant is able to retrieve content from websites such as YouTube, Wikipedia, and Chrome. For instance, it can pull data from Google or Wikipedia or ask for a particular video to play. Initially, the assistant asks the user what they want it to do. The assistant utilizes the element's XPath to locate the search bar and clicks on it when the user wants a search and the browser opens. The assistant then types the user's desired search term. Similar to how it did with the search bar, the assistant uses its XPath to click the search button after submitting the request. This is how the browser and assistant work together.  
The assistant searches for the video and hits the search button after opening YouTube in the browser when a user requests that it play a YouTube video. After that, it uses its XPath to find the video and selects the first one found in the search results.

COMMUNICATION INTERFACE MODULES

Last but not least, we used the serial module to integrate IoT capability into the project. In order to operate and communicate with Internet of Things devices, this module opens the serial interface of the Arduino board via ports 11 and COM3.

*Algorithm:*

* *Automatic speech recognition (ASR)*

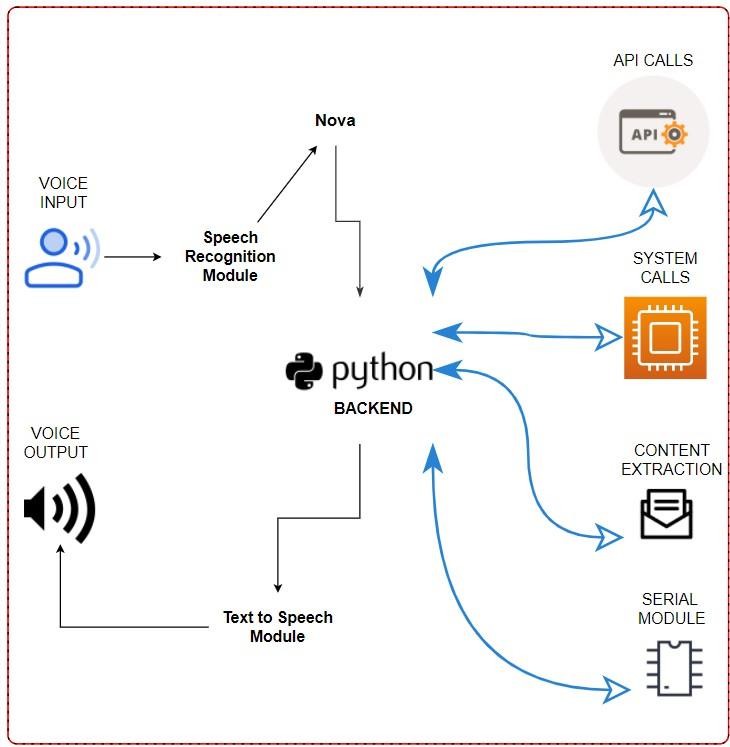
This system makes use of the Recognizer class, which makes it easier to turn audio files into text before producing voice output. Values below this threshold are regarded as silence, while values over it are seen as communication. The energy threshold function establishes the energy level for sounds. In order to ensure optimal speech recognition, the recognizer instance dynamically modifies the energy threshold based on ambient noise, adjusting it according to the source and duration of the noise.

* *Conversion from Speech to Text and Text to Speech*

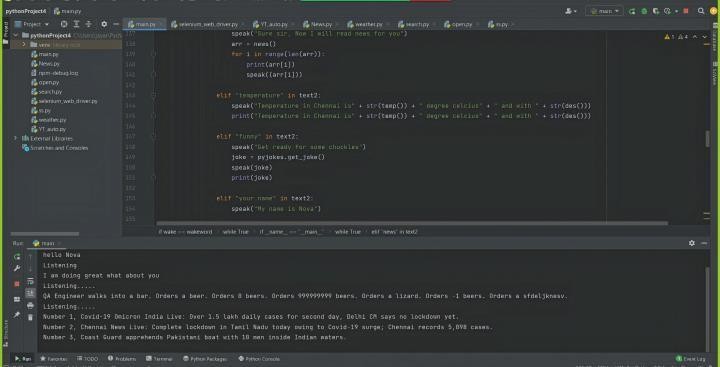
Pyttsx3 is a Python text-to-speech conversion tool. It is able to alter the voice, rate, and volume with particular orders. To convert audio to text for additional processing, Python has an API called Speech Recognition. This allows us to turn lengthy or huge audio recordings into text. We have included the eSpeak and SAPI5 TTS engines, which are capable of processing the same.

*-Process & Executes the Required Command*

In order to execute these commands, the aforementioned command is first translated into text using a voice recognition module, then it is placed in a temporary memory. Next, it analyzes the user's text in a temporary memory,the user's input is used to determine what the system needs, and a while loop is started to process the user's commands continually until a certain condition is fulfilled or the work is finished.



highlighting its expanding capabilities. As it develops, we can anticipate that the voice assistant will rank among the most important innovations in the current technological environment.

Although there has been some discussion over additional development, the program is nearly finished on our end and functioning as planned. We expect to make advancements soon that will improve the helper we have developed and make it much more beneficial what its current state.

**Fig 4. Voice Assistant System**

1. CONCLUSION

As mentioned earlier, In the contemporary world, "the voice assistant is one of the biggest problem solvers," as the instances given make abundantly evident. It's clear from looking at the suggested features that voice assistants are quickly becoming one of the biggest developments in artificial intelligence to date. Voice assistants used to be limited to simple functions like stating the date or conducting web searches. However, it's evident that voice assistants have evolved into extremely sophisticated software given their current extended capabilities. The objective of this project is to make the assistant even more advanced by incorporating more intelligent applications. Users will be able to save a significant amount of time as a result. To sum up, we're dedicated to creating one of the greatest voice assistants out there.

1. REFERENCES

**Fig 3. Architecture Diagram**

1. Results

This study provides a concise synopsis of the voice assistant project, emphasizing its benefits and the manner in which it may help users accomplish a range of activities. It also talks about the assistant's ongoing development,

Easwara Moorthy, Aarthi & Vu, Kim-Phuong, “Voice Activated Personal Assistant: Acceptability of Use in the Public Space” HIMI 2014. Lecture Notes in Computer Science, vol 8522. Springer, pp. 324-334, 10.1007/978- 3-319-07863-2\_32.

V. Këpuska and G. Bohouta, "Next-generation of virtual personal assistants (Microsoft Cortana, Apple Siri, Amazon Alexa and Google Home)," 2018 IEEE 8th Annual Computing and Communication Workshop and Conference (CCWC), 2018, pp. 99-103, doi:

10.1109/CCWC.2018.8301638.

George Terzopoulos and Maya Satratzemi. 2019. “Voice Assistants and Artificial Intelligence in Education”. In Proceedings of the 9th Balkan Conference on Informatics (BCI'19). Association for Computing Machinery, New York, NY, USA, Article 34, 1–6. DOI:https://doi.org/10.1145/3351556.3351588

R. Kumar, G. Sarupria, V. Panwala, S. Shah and N. Shah, "Power Efficient Smart Home with Voice Assistant," 2020 11th International Conference on Computing, Communication and Networking Technologies (ICCCNT), 2020, pp. 1-5, doi: 10.1109/ICCCNT49239.2020.9225612.

T. -K. Kim, "Short Research on Voice Control System Based on Artificial Intelligence Assistant," 2020 International Conference on Electronics, Information, and Communication (ICEIC), 2020, pp. 1-2, doi: 10.1109/ICEIC49074.2020.9051160.

S. Subhash, P. N. Srivatsa, S. Siddesh, A. Ullas and B. Santhosh, "Artificial Intelligence-based Voice Assistant," 2020 Fourth World Conference on Smart Trends in Systems, Security and Sustainability (WorldS4), 2020, pp. 593-596, doi: 10.1109/WorldS450073.2020.9210344.

Dellaert, B.G.C., Shu, S.B., Arentze, T.A. et al. Consumer decisions with artificially intelligent voice assistants. Mark Lett 31, 335–347 (2020). https://doi.org/10.1007/s11002-020-09537-5

Harshit Agrawal, Nivedita Singh, Gaurav Kumar, Dr. Diwakar Yagyasen, Mr. Surya Vikram Singh “Voice Assistant Using Python” 2021, IJIRT Volume 8 Issue 2, ISSN: 2349-6002, pp.419-423.

Abeed Sayyed, AshpakShaikh, AshishSancheti,Swikar Sangamnere, Prof. Jayant H Bhangale. “Desktop Assistant AI Using Python” (2021) International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)Volume 6, Issue 2, June 2021. ISSN (Online) 2581-942.

P.Krishnaraj,F.Mohamed Faris,D.Rajesh “Portable Voice Recognition with GUI Automation” 2021, IJIRT, Volume 9 Issue 6 ISSN (Print): 2320-9356 PP.

20-23

Mrs.A.M.Sermakani, J.Monisha, G.Shrisha, G.Sumisha, “Creating Desktop Speech Recognization Using Python Programming.” IJARCCE, Vol. 10, Issue 3, March 2021, ISSN (Online), pp.129-134.