IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.67
Volume 5, Issue 2, March 2025

Amigo Voice Assistant

Ms. Tanvi Mahale[1], Ms. Sakshi Mane[2], Ms. Yashaswi Walunj[3], Mrs. Vijaya Chavan[4]
Students, Department of Computer Technology[1,2,3]
Lecturer, Department of Computer Technology[4]
Bharati Vidyapeeth Institute of Technology, Navi Mumbai, Maharashtra, India.

Abstract: *A Voice Assistant is one of the hot topics in the current world that are programs that listens to human's verbal command and respond to them which makes it a human computer/device interaction. In the current days, a voice assistant is everywhere which is a lot useful in these busy days. Nowadays, almost everyone in the current world is using voice assistant because it's everywhere starting from Google smartphone assistant which even 5 years old kids will know how to use because of the current world pandemic which makes them use smartphones till Amazon's Alexa which will be very useful to do works starting from entertaining the users till turning on and off the household products (Internet of Things). One of the greatest features is that it will be very useful to even physically challenged people, for example, people who aren't able to walk use the Internet of Things (IoT) feature to operate the household products and maintain them. So, we tend to develop a voice assistant which will be very useful to the users same as the other voice assistants which are currently in the world. A Voice Assistant is one of the hot topics in the current world that are programs that listens to human's verbal command and respond to them which makes it a human computer/device interaction. In the current days, a voice assistant is everywhere which is a lot useful in these busy days. Nowadays, almost everyone in the current world is using voice assistant because it's everywhere starting from Google smartphone assistant which even 5 years old kids will know how to use because of the current world pandemic which makes them use smartphones till Amigo assistant. which will be very useful to do works starting from entertaining the users till turning on and off the household products (Internet of Things). One of the greatest features is that it will be very useful to even physically challenged people, for example, people who aren't able to walk use the Internet of Things (IoT) feature to operate the household products and maintain them. So, we tend to develop a voice assistant which will be very useful to the users same as the other voice assistants which are currently in the world. A Voice Assistant is one of the hot topics in the current world that are programs that listens to human's verbal command and respond to them which makes it a humancomputer/device interaction. In the current days, a voice assistant is everywhere which is a lot useful in these busy days. Nowadays, almost everyone in the current world is using voice assistant because it's everywhere starting from Google smartphone assistant which even 5 years old kids will know how to use because of the current world pandemic which makes them use smartphones till Amazon's Alexa which will be very useful to do works starting from entertaining the users till turning on and off the household products (Internet of Things). One of the greatest features is that it will be very useful to even physically challenged people, for example, people who aren't able to walk use the Internet of Things (IoT) feature to operate the household products and maintain them. So, we tend to develop a voice assistant which will be very useful to the users same as the other voice assistants which are currently in the world*

Keywords: Voice Assistant

I. INTRODUCTION

In today's fast-paced digital world, voice assistants have become an integral part of modern technology. These AI-powered tools offer hands-free interaction with devices, making it easier for users to perform tasks, access information, and streamline daily activities. The aim of this project is to develop a voice assistant that can understand natural language commands and respond accordingly.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.67
Volume 5, Issue 2, March 2025

The voice assistant will leverage advanced speech recognition and natural language processing (NLP) technologies to interpret user commands, enabling users to interact with their devices in a more intuitive way. Whether it's setting reminders, controlling smart devices, answering questions, or playing music, the voice assistant aims to simplify complex processes and improve the overall user experience.

II. METHODOLOGY

PROPOSED SYSTEM

We are proposing a system in an efficient way of implementing a Personal voice assistant, Speech Recognition library has many in-built functions, that will let the assistant understand the command given by user and the response will be sent back to user in voice, with Text to Speech functions. When assistant captures the voice command given by user, the under lying algorithms will convert the voice into text. And according to the keywords present in the text (command given by user), respective action will be performed by the assistant.

This is made possible with the functions present in different libraries. Also, the assistant was able to achieve all the functionalities with help of some API's. We had used these APIs for functionalities like performing calculations, extracting news from web sources, and for telling the weather. We will be sending a request, and through the API, we're getting the respective output. API's like WOLFRAMALPHA, are very helpful in performing things like calculations, making small web searches. And for getting the data from web. In this way, we are able to extract news from the web sources, and send them as input to a function for further purposes. Also, we have libraries like Random and many other libraries, each corresponding to a different technology. We used the library OS to implement Operating System related functionalities like Shutting down a system, or restarting a system.

At the outset we make our program capable of using system voice with the help of sapi5 and pyttsx3. pyttsx3 is a text-to-speech conversion library in Python. Unlike alternative libraries, it works offline, and is compatible with both Python 2 and 3. The Speech Application Programming Interface or SAPI is an API developed by Microsoft to allow the use of speech recognition and speech synthesis within Windows applications. Then we define the speak function to enable the program to speak the outputs.

After that we will define a function to take voice commands using the system microphone. The main function is then defined where all the capabilities of the program are defined.

The proposed system will have the following functionality:

1) The system will keep listening for commands and the time for listening is variable which can be changed according to user requirements

2) If the system is not able to gather information from the user input it will keep asking again to repeat till the desired number of times.

3) The system can have both male and female voices according to user requirements.

4) Features supported in the current version include playing music, texts, search on Wikipedia, or opening system installed applications, opening anything on the web browser, etc.

III. DISCUSSION

The development of a voice assistant represents a fascinating intersection of various fields such as artificial intelligence (AI), natural language processing (NLP), machine learning (ML), and human-computer interaction (HCI). As we progress in this voice assistant project, several key aspects need to be considered, from technical challenges to user experience. In this discussion, we will explore the various components of the project, challenges faced, and the potential impact of the voice assistant in real-world applications.

1. Core Technologies Involved

The voice assistant relies heavily on several core technologies:

- **Speech Recognition:** The first step in any voice assistant is the conversion of spoken words into text. This is accomplished through speech recognition algorithms. Common APIs like Google Speech-to-Text or Microsoft Azure Speech API can be used, but building a custom model using frameworks like Kaldi or DeepSpeech is also possible for specific needs.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.67
Volume 5, Issue 2, March 2025

- **Natural Language Processing (NLP):** After converting speech to text, the voice assistant needs to understand the meaning behind the commands. NLP is essential for this, as it enables the assistant to interpret and process the text in a human-like way. NLP tasks such as tokenization, named entity recognition (NER), and syntactic parsing help in understanding complex queries.
- **Text-to-Speech (TTS):** Once the assistant processes a command, it needs to respond to the user. TTS technology converts the assistant's response from text into speech. Services like Google TTS and Amazon Polly are commonly used for this purpose, but custom voice synthesis solutions can also be designed to add personality and uniqueness to the assistant.

2. Key Features of the Voice Assistant

- **Speech Command Understanding:** The core functionality of any voice assistant is its ability to interpret user commands. Simple tasks like setting reminders, controlling home automation systems, and searching the web are standard, but more complex requests such as conversational context and understanding multi-turn dialogues are essential for a superior experience.
- **Contextual Awareness:** A good voice assistant should not only respond to individual commands but also understand the context of previous interactions. This can involve remembering the user's preferences, habits, or location, and using this information to deliver more personalized responses.
- **Multimodal Interaction:** While voice assistants are typically focused on voice interactions, the ability to integrate visual or touch-based feedback (such as showing results on a screen) can enhance the user experience, especially for tasks like weather updates, navigation, or multimedia control.
- **Integration with Third-Party Services:** The effectiveness of a voice assistant depends on how well it integrates with external services, such as music streaming platforms (Spotify, Apple Music), smart home devices (Philips Hue, Nest), calendars, emails, etc. API integrations make this possible, allowing the assistant to control devices, schedule appointments, or answer questions by pulling data from external sources.

3. Challenges Faced

- **Accuracy in Speech Recognition:** One of the most significant challenges is achieving high accuracy in speech recognition, especially in noisy environments or with diverse accents and dialects. This challenge can be mitigated by using advanced speech models trained on large, diverse datasets.
- **Language and Context Understanding:** Voice assistants must handle a wide range of languages and dialects, as well as understand the context behind commands. Developing the system to understand nuances like humor, ambiguity, and multi-turn conversations remains an ongoing research challenge.
- **Privacy and Security:** Voice assistants are designed to listen to and process user conversations, raising concerns about privacy. Ensuring that users' data is securely stored and only accessed when needed is a crucial aspect of the project. Implementing strong encryption and offering users control over their data (e.g., the ability to delete voice recordings) is critical.
- **Real-Time Response and Processing:** In real-world scenarios, the system needs to process commands and provide responses in real time. This requires low-latency systems that efficiently handle speech input, process it, and deliver output without noticeable delay.
- **Hardware Constraints:** While cloud-based systems can handle much of the heavy lifting, voice assistants embedded in devices (e.g., smartphones, smart speakers) must often operate with limited computational power. This makes it challenging to deploy complex machine learning models without affecting performance.

4. Potential Applications and Impact

Voice assistants have numerous applications in daily life, including:

- **Smart Homes:** Voice assistants can control home automation systems, such as lights, thermostats, locks, and entertainment systems, improving convenience and accessibility.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.67
Volume 5, Issue 2, March 2025

- **Healthcare:** In healthcare, voice assistants can assist patients by setting medication reminders, providing health tips, or even helping medical professionals with hands-free tasks during procedures.
- **Business and Productivity:** Voice assistants can schedule meetings, manage emails, or automate repetitive tasks, improving productivity in both personal and professional environments.
- **Education:** Voice assistants can assist with language learning, provide quick answers to academic questions, or even act as interactive tutors.

IV. WORKING OF THE PROJECT

- It starts with a signal word. Users say the names of their voice assistants for reason.
- They might say, “Hey Siri!” or simply, “Amigo!” Whatever the signal word is, it wakes up the device.
- It signals to the voice assistant that it should begin paying attention. After the voice assistant hears its signal word, it starts to listen.
- The device waits for a pause to know you’ve finished your request. The voice assistant then sends our request over to its source code.
- Once in the source code, our request is compared to other requests. It’s split into separate commands that our voice assistant can understand.
- The source code then sends these commands back to the voice assistant. Once it receives the commands, the voice assistant knows what to do next.
- If it understands, the voice assistant will carry out the task we asked for. For example, “Hey AMIGO! What’s the weather?” AMIGO reports back to us in seconds.
- The more directions the devices receive, the better and faster they get at fulfilling our requests.
- The user gives the voice input through microphone and the assistant is triggered by the wake up word and performs the STT (Speech to Text) and converts it into a text and understands the Voice input and further performs the task said by the user repeatedly and delivers it via TTS (Text to Speech) module via AI Voice.

V. LITERATURE SURVEY

Related work

This field of virtual assistants having speech recognition has seen some major advancements or innovations. This is mainly because of its demand in devices like smartwatches or fitness bands, speakers, Bluetooth earphones, mobile phones, laptop or desktop, television, etc. Almost all the digital devices which are coming nowadays are coming with voice assistants which help to control the device with speech recognition only. A new set of techniques is being developed constantly to improve the performance of voice automated search.

As the amount of data is increasing exponentially now known as Big Data the best way to improve the results of virtual assistants is to incorporate our assistants with machine learning and train our devices according to their uses. Other major techniques that are equally important are Artificial Intelligence, Internet of Things, Big Data access and management, etc. With the use of voice assistants, we can automate the task easily, just give the input to the machine in the speech form and all the tasks will be done by it from converting your speech into text form to taking out keywords from that text and execute the query to give results to the user.

Machine Learning is just a subset of Artificial Intelligence. This has been one of the most helpful advancements in technology. Before AI we were the ones who were upgrading technology to do a task but now the machine is itself able to counter new tasks and solve it without need to involve the humans to evolve it.

VI. FUTURE SCOPE

We are entering the era of implementing voice-activated technologies to remain relevant and competitive. Voice-activation technology is vital not only for businesses to stay relevant with their target customers, but also for internal operations. Technology may be utilized to automate human operations, saving time for everyone. Routine operations, such as sending basic emails or scheduling appointments, can be completed more quickly, with less effort, and without the use of a computer, just by employing a simple voice command. People can multitask as a result, enhancing their

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.67
Volume 5, Issue 2, March 2025

productivity. Furthermore, relieving employees from hours of tedious administrative tasks allows them to devote more time to strategy meetings, brainstorming sessions, and other jobs that need creativity and human interaction.

1) Sending Emails with a voice assistant: Emails, as we all know, are very crucial for communication because they can be used for any professional contact, and the finest service for sending and receiving emails is, as we all know, GMAIL. Gmail is a Google-created free email service. Gmail can be accessed over the web or using third-party apps that use the POP or IMAP protocols to synchronize email content.

To integrate Gmail with Voice Assistant we have to utilize Gmail API. The Gmail API allows you to access and control threads, messages, and labels in your Gmail mailbox

2) Scheduling appointments using a voice assistant: The demands on our time increase as our company grows. A growing number of people want to meet with us. We have a growing number of people wh0 rely on us. We must check in on certain projects or set aside time to chat with possible business leads. There won't be enough hours in the day if we keep doing things the old way

We need to get a better handle on our full-time schedule and devise a strategy for arranging appointments that doesn't interfere with our most critical job. By working with a virtual scheduler or, in other words, a virtual assistant, we let someone else worry about the organization and prioritize our schedule while we focus on the work.

VII. CONCLUSION

Voice assistants have evolved from simple command-based systems to sophisticated tools that seamlessly integrate into our personal and professional lives. Their key features—ranging from speech recognition and personalization to multi-platform integration—have made them indispensable in modern society.

As technology advances, voice assistants are becoming more intelligent, context-aware, and capable of handling complex tasks. They are transforming industries by enhancing customer support, improving productivity, and providing personalized experiences.

In conclusion, voice assistants are not just a passing trend but a fundamental shift in how we interact with technology. Their continuous development promises to further enrich our daily lives, making interactions more natural, efficient, and accessible.

As stated before, "voice assistant is one of the biggest problem solver" and you can see that in the proposals with the examples that it is in fact one of the biggest problem solver of the current world.

The main idea is to develop the assistant even more advanced than it is now and make it the best ai in the world which will save an ample of time for its users

REFERENCES

[1]. K. Noda, H. Arie, Y. Suga, T. Ogata, Multimodal integration learning of robot behavior using deep neural networks, Elsevier: Robotics and Autonomous Systems, 2014.

[2]. Artificial intelligence (AI), sometimes called machine intelligence. https://en.wikipedia.org/wiki/Artificial_intelligence.

[3]. Deepak Shende, RiaUmahiya, Monika Raghorte, AishwaryaBhisikar, AnupBhange, "AI Based Voice Assistant Using Python", Journal of Emerging Technologies and Innovative Research (JETIR), February 2019, Volume 6, Issue 2.

[4]. J. B. Allen, "From lord rayleigh to shannon: How do humans decode speech," in International Conference on Acoustics, Speech and Signal Processing, 2002.

[5]. John Levis and Ruslan Suvorov, "Automatic Speech Recognition

[6]. B.H. Juang and Lawrence R. Rabiner, "Automatic Speech Recognition - A Brief History of the Technology Development