Personal Voice Assistant
Using Computer Vision

1ˢᵗDr A.Jagan
Computer Science Of Engineering(of JNTU Aff.)
B V Raju Institute Of Techology (of JNTU Aff.)
Narsapur,Medak,Telangana,India
jagan.amgoth@bvrit.ac.in

2ⁿᵈD.Parvathi
Computer Science Of Engineering (of JNTU Aff.)
B V Raju Institute Of Techology (of JNTU Aff.))
Narsapur,Medak,Telangana,India
21211A0572@bvrit.ac.in

3ʳᵈH.Susanna
Computer Science Of Engineering (of JNTU Aff.)
B V Raju Institute Of Techology (of JNTU Aff.)
Narsapur,Medak,Telangana,India
21211A05A6@bvrit.ac.in

4ᵗʰD.Nandini
Computer Science Of Engineering(of JNTU Aff.)
B V Raju Institute Of Techology (of JNTU Aff.)
Narsapur,Medak,Telangana,India
21211A0575@bvrit.ac.in

5ᵗʰMohammad Ghouse Pasha
Computer Science Of Engineering(of JNTU Aff.)
B V Raju Institute Of Technology(of JNTU Aff.)
Narsapur,Medak,Telangana,India
22215A0508@bvrit.ac.in

Abstract—Modern society has made daily life smarter and more technologically connected to make it easier to go about our daily tasks. A voice assistant converts speech to text ,then process the text, convert text to speech and add various functionalities. In the era of smart personal assistants, Amazon Alexa, Apple Siri, and few other voice assistants have found popularity. A voice assistant interacts with humans by listening to their verbal commands and acting upon them. "SIRI," the most well-known iPhone app, allows voice communication between users' mobile devices and responds to their voice commands. In addition, Google developed a similar technology called "Google Voice Search," which is used on Android smartphones. However, these applications mostly utilizes internet connections. Our proposed model works without internet connections and can manage tasks efficiently.

Key words: Voice Assistant, Scheduling, Text, Speech, Reminders

I. INTRODUCTION

Voice Assistant is an application that carries out everyday tasks via voice commands. It combines AI and machine learning to recognize our speech and carry out our requests. Voice assistants are gadgets or applications that reply to human speech using AI, voice recognition, and NLP. The technology allows the gadget to synthesize, deconstruct, assess, and provide a relevant response to the user's message. Since technologies allow users to connect with their devices and access information conveniently and hands-free, personal voice assistants have grown in popularity in recent years. In addition to playing music and setting alarms, personal voice assistants may also answer queries and make phone calls. One area where personal voice assistants have the potential to be particularly useful in scheduling and time management.

Personal voice assistants as schedulers allow users to manage their calendars, schedule appointments, set reminders, and create to-do lists, all through natural language voice commands. This hands-free approach to scheduling can help users save time and stay organized, without the need to manually enter information into a calendar or reminder app.

Additionally, personal voice assistants can provide personalized recommendations based on the user's behavior and preferences. For example, a personal voice assistant might suggest a time for a meeting based on the user's availability and location, or provide reminders to take a break if the user has been working for an extended period of time.

Overall, personal voice assistants as schedulers have the potential to improve time management and productivity, while providing a convenient and hands-free user experience. With advances in natural language processing and machine learning, personal voice assistants are becoming increasingly accurate and responsive, making them a promising tool for scheduling and time management in the future.

II. LITERATURE SURVEY

Today, we train our robots to act like people and do activities on their own, taking the place of humans with machines. Based on this circumstance, the idea of a voice assistant emerges, capable of doing a variety of tasks for people based just on their speech. The virtual assistant may filter out certain user instructions and deliver information that is relevant to the command.

The AIVA (Cortana from Microsoft, Google Assistant from Google, and the newly debuted intelligent assistant under the

name "AIVA" 2018), according to Deepak Shende and Ria Umabiya, aimed at producing a voice-controlled personal assistant that is performing many tasks, such searching the Internet. It contains several brand-new features, such the ability to submit comments on social networking platforms like Facebook, Twitter, etc. using a few simple instructions. Additionally, you may get the local weather information and the local climate in your area. Tulshan clarified that the user's fingers may suffer damage as a result of constant typing. To prevent these issues, we must create a system that enables us to do tasks using voice commands. The system will detect the voice, synthesis the recognized words, and, if they are acceptable or make sense, display them on the screen. The program will then be assembled and run by the system once again by identifying the precise keywords[1].

Speech synthesis might be a method for creating human speech in an artificial environment. The output of speech synthesis is an audio stream. On the other hand, a speech recognizer does the opposite task. It converts an audio stream into text by taking an input of an audio stream. A voice synthesizer is a kind of computer program that does speech synthesis; it is often employed in hardware or software devices. Combining the segments of recorded speech that are kept in a database is often used to produce synthesized speech. Naturalness and understandability are two of a voice synthesizer's most crucial characteristics. Naturalness is the extent to which the output is similar to human speech, while intelligibility is the ease with which the output may be understood. The ideal speech synthesizer is both natural and understandable[2].The voice recognition model is one of the most essential components of a virtual assistant. Given the multiple Neural Networks required to create a speech recognition system, it was crucial to evaluate the models that provided the insight by examining their accuracy and other characteristics. When compared to a basic neural network, it was shown that the convolutional neural network (CNN) model obtained High Accuracy and Lower Validation Accuracy. demonstrating why CNN is a superior option for voice recognition systems. In order to scale voice recognition using CNN, the following factors were considered:

(i) Throughput,

(ii) Real-Time Factor (RTF) and latency(iii) Word

Error Rate (WER)[3].

In order to perform voice search, a two-stage search procedure is used, in which string candidates generated by an automatic speech recognition (ASR) system are rescored in order to locate the best matching entry from a potentially very large application-specific database.[4]

A domain agnostic ASR system may be utilized in conjunction with additional domain specific knowledge sources, as shown in Study, to provide voice access to internet search indexes. The logical method to enhance recognition is to provide additional data to a particular voice recognition problem, according to International Journal of Engineering Research Technology. Microsoft created Cortona as a personal virtual assistant for Windows, iOS, Android, and other platforms. Only Windows 10 supports Cortona in the Windows operating system. In 2015, it was made available for Windows 10. For usage with the Cortona program we attempt to configure for activating the Cortona in our laptops or PCs, Cortona is in icon form on the taskbar next to the search bar in Windows 10. Although setup takes more time, searching is simple. It takes a lot of time and only Windows 10 is compatible with it [5].

Personal and virtual assistants are becoming an essential part of our everyday life. All organizations and people are adopting these technologies since they make it easier for them to do their responsibilities. This system is based on a desktop application. This system includes a virtual assistant that can take input from users, understand it, evaluate it, and do tasks as needed. Users may save a ton of time by doing this[6].

The new model of VPAs will be used to encourage interaction between humans and machines by using a number of technologies, including gesture detection, image/video recognition, detection of voices, a significant dialogue and interactive knowledge base, and an extensive knowledge base. Other potential applications for the new VPAs system include automation of homes, security access control, robotics and vehicles, solutions for the handicapped, and education.[7].A virtual voice-based personal intelligent assistant for visually impaired people was introduced by Aditya Sinha et al. Virtual Personal Assistants (VPAs) are the next generation carrier services for mobile and smart device customers, according to Moustafa Elshafei. VPA offers a single point of contact that seamlessly engages a variety of information and efficiently reacts to conversational voice commands. Additionally, it organizes calls, calendars personal activities, allows users to access task managers through voice interface, and has all the features of unified messaging.[8] The virtual personal assistant helps the user to reduce disruptions to routine workflow, maximize time and expense, and increase overall productivity. There have been several surveys done in the field of voice recognition. For instance, Morgan reviewed feed-forward networks that have been discriminatively trained to help in voice recognition. The major goal of the study was to highlight works that use numerous levels of processing before decoding word sequences using a hidden Markov model. In the end, this review study came to the conclusion that, even while deep processing structures are capable of improving this genre, layer

thickness and feature selection may also be important considerations[9].

A speech controlled automatic wheelchair that functions as a smart home navigation system was suggested by M. Senthil Sivakumar. It enables elderly or physically challenged people to navigate about their homes using voice command. The wheelchair also has an ultrasonic sensor-based obstacle avoidance system. Megha Muralidharan presented an intelligent wheelchair powered by speech recognition, in which a voice recognition module recognizes the patient's voice to operate the wheelchair. By utilizing their voice, the user may go forward, backward, right, and left. The system uses hand gestures to control the joystick in addition to two ultrasonic sensors to detect the depth of barriers. B. Gokul suggests Android-based closed-loop speed control of a DC motor using Bluetooth voice recognition. Through a smartphone app, this technology provides speech recognition. This program receives the patient's speech, which is subsequently sent over Bluetooth to the Arduino. The motor driver IC is then instructed to operate as necessary via Arduino. The advantage of utilizing a speed sensor to determine the motor speed is another benefit of this technology. In the app, speed information is presented. It was recommended to utilize the Raspberry Pi as the processing unit rather of an Arduino or microcontroller in a new system that Romil Chauhan has not yet created, dubbed Study on Implementation of Voice Controlled Wheelchair. The Raspberry Pi may be linked to an obstacle detection sensor. It will lessen the need for hardware like motor drivers. It is simply a suggested proposal, though[10].

A variety of techniques for improving the effectiveness of automated voice search for mobile users accessing these services via a range of portable devices. Voice search is implemented using a two-stage search process in which string candidates generated by an automatic speech recognition (ASR) system are rescored in order to discover the best matching entry from a database that is specifically designed for the application in question. A domain agnostic ASR system may be utilized in conjunction with additional domain specific knowledge sources, as shown by Study, to provide voice access to internet search indices[11].

III. Problem Statement

- To create a personal voice assistant for user to manage tasks and schedules effectively .
- The proposed system will be more accurate than previously existing systems.
- Previous systems doesn't work in online mode and they need registered accounts.

IV. Objectives

1) A model that helps users to stay organized by reducing risk of missing important events.
2) This model helps users to input and view their schedules.
3) The users can edit their schedules to manage their tasks.

V. Existing Work

Some of the existing works on voice assistants are Apple Siri, Microsoft's Cortana and few which uses language processing and speech recognition concepts. They pay attention to user needs and successfully complete every task given to them.

Google Assistant: Virtual assistant Google Assistant, which mostly works on mobile and smart home devices, makes use of Artificial Intelligence. Google has a reputation for being user-friendly. You can access it via your speaker, watch, laptop, TV, or automobile. It is capable of establishing connections with widely used products and programs including Netflix, YouTube, and Spotify. The same output formats are available for both voice and text input.

Microsoft's Cortana: In essence, Cortana is a digital assistant that can learn from the user's requests and then adapt its behaviour when the user uses it again. It can also accept text or voice inputs. Additionally, it offers comparable services like calendar management, virtual assistant chat, search functionality, etc.

Amazon's Alexa: Alexa was developed by Amazon, who did an extremely good job. This virtual assistant has the capacity to integrate various technologies into our daily lives. Due to its ability to link a vast number of devices, including lights, televisions, and numerous other electrical appliances, Alexa is particularly well-liked in home automation systems. It may also connect you to the newest songs on the market and monitor the weather and news for you. Additionally, using voice commands, it can shop for the user and reserve a cab.

However, in order to utilize these assistants, you must have an account (such as a Google or Microsoft account), as well as a wireless connection, since they can only be used while they are linked to the internet. They work with a variety of devices, including laptops and mobile phones.

VI. Proposed Work

To provide a user interface that enables digital device control without the need of a hand. It will analyze the possible use of one particular piece of software as a VAs by looking at

illustrations of intelligent software using NLP that are already available, with different types of guidance. The procedure starts with the user's audio commands being analyzed using the microphone.

This might include receiving any information, such as a schedule, remaining time, or a timetable by an audio command. The concept of a voice assistant may soon become a reality, according to certain emerging technology. Reading the sources mentioned above and putting their examples to the test served as a base for my study. In order to learn more about voice assistants and best practices, tests are created by programming in accordance with books and internet resources.

User testing and experiments on this system have shown that a simple software utilizing Python algorithms may function as a digital personal assistant. Depending on the user's high-quality input, iteratively learn the user's preference for each heme using Python algorithms. voice assistant as a digital service, in idea Finding solutions for all of our requirements is quickly becoming a reality. Digital and voice assistants are poised to replace us as internet users and become more familiar with us than we are ourselves as artificial intelligence and machine learning advance quickly. Google Now and Siri are just the beginning. The system receives voice input and processes it using a variety of Python and machine learning.

VII. METHODS

As we all are aware that every human has and unique characteristics and appiles his own methods for the product development. However, most assistants have the capability to complete a more limited set of tasks but with greater accuracy and in accordance with the user's preferences. There is no assistant who can perform the tasks equally as an outcome. The qualities of an assistant are determined by the field on which the developer concentrates.

The basic idea behind creating a voice assistant remains the same, regardless of the various methods used to learn various algorithms. The technologies employed are Speech recognition, Text-to-Speech, dialogue managers, natural language comprehension, and named entity recognition which can all be used to create a voice assistant that can communicate with people.

figures/voice.jpg

Fig. 1. Technologies used for constructing Intelligent systems of interaction with a human by natural Language

A. Working Process

(a) The system will pay attention to the user's orders and instructions, and the user can alter the listening time based on their demands.
(b)The system will constantly ask the user to repeat till a certain number of times if it is unable to get data from user input.
(c) According to user requirements, the system can feature both male and female voices.
(d) Timetable-based features, a Schedule reader, a Schedule reminder, etc.

B. System Architecture

figures/result3.jpg

Fig. 2. The Working Of the Voice Assistant

Speech Recognition : You'll need a mechanism to translate spoken words into text, which is speech recognition. In Python, the SpeechRecognition library is an appreciated option.

It features a user interface for a number of voice recognition APIs, including Sphinx and Google voice Recognition.

Pyttsx3: Pyttsx3 is a widely used Text-to-Speech library that allows user to convert text into speech in a consistent manner by providing desired text as input. Pyttsx3 provides straightforward and easy-to-use interface.

Speech Synthesis: Speech synthesis, also referred to as text-to-speech (TTS), is the process of converting text into spoken words. It enables you to produce speech from text, which can be beneficial for a number of applications, including voice assistants, accessibility tools, and the creation of audio content. In Python, there are several libraries available that can help you generate speech output from text. To provide spoken responses, you'll need to convert text back into speech. Libraries like pyttsx3, gTTS (Speech Synthesis/Google Textto-Speech), and espeak can generate speech output from text.

Command Processing: Based on the user's input and recognized intent, determine the appropriate action to be taken. If the user wants to add an event, extract the necessary details and create a new event object.

Date-Time Module: The date and time module handles processing and manipulation of date and time related information to perform tasks like validating time and date inputs. This module allows user to format a date or time into a string.

Schedule Module: The schedule module allows users to schedule and manage tasks at specific time and runs the function at a specific time.

figures/result4.jpg

Fig. 3. Module to Interpret Voice Response Text-to-Speech

VIII. RESULTS

Users get various advantages from using voice assistants. It takes less time with a virtual voice assistant. Software that can do tasks assigned to it by an end user and understand verbal or written commands is known as a virtual voice assistant. Virtual assistants employ speech synthesis to convert user voice or text input into effective guidance. With the help of a virtual voice assistant, you can manage laptops and desktops on your own schedule. Time is saved since it is a rapid procedure. Because they are working for the user during certain hours, virtual voice assistant is accessible to the user at all times and can swiftly adjust to changing demands. You will have access to a virtual voice assistant, that, if their workload permits it, can also help friends, family, and colleagues.

1) *Efficient Scheduling*: Through the voice assistant, users can rapidly create, change, and manage their events with voice commands, streamlining the scheduling process.
2) *Hands-Free Operation*: Without the need for manual input, the user can communicate with the personal assistant simply using voice commands, offering a hands-free experience.
3) *Timetable Retrieval*: By asking the assistant, users may quickly get their daily or weekly schedules. The schedule information is spoken aloud by the voice assistant, enabling users to immediately organize their activities.
4) *Reminders*: Based on user requests, the assistant can provide reminders for certain times or activities, assisting users in staying organized and on track.
5) *Time and Productivity Management*: By giving consumers quick access to their calendars, reminders, and deadlines, the voice assistant enables better time management.

As the voice Assistant has the following benefits the outcome will be as shown:

1)Timetable will be displayed.

2)Add reminders to the schedule. Overall, our personal voice assistant helps the users to stay organized about their schedules and reminds the users about their tasks.

Fig. 4. Result

IX. DISCUSSIONS

The Python backend handles the fundamental logic and functionality of a personal voice assistant for scheduling. It accepts user requests or instructions, evaluates them, and then performs the necessary tasks related to task scheduling.

Our main objective is to enable people do their jobs using voice commands. You can accomplish this in two steps. Using speech recognition, first take the user's audio input and translate it into an English sentence. The second is to track down the intended user of the activity and send it to the server. via the HTTP Protocol, and display the outcome. The initial outcome that gets generated after the Windows Code has been run will start listening. Following the user's voice command When a user gives a voice command and translates it into an English phrase, The user's command will be shown on your device upon translation.

X. CONCLUSION

The following document provides a thorough description of a Voice Activated Personal Assistant (VAPA) developed in Python that allows users to check their calendars by speaking instructions, reducing the need for manual input and accelerating the scheduling process.Building a voice assistant using Python can be a rewarding endeavor. Python provides a rich ecosystem of libraries, frameworks and tools that can assist in developing various components of a voice assistant, such as speech recognition and response generation. Our personal voice assistant is able to do almost everything that From opening a specific schedule to reminding them of the day's task in the timetable, the user gives it instructions. Using Natural Language Processing, we kept our approach to the challenge straightforward.Most of the user's activities are made easier by it, including searching timetables and schedules. Overall, Personal voice assistant have the potential to greatly enhance our daily lives

XI. REFERENCES

[1]Pooja C. Goutam, Monika S.Jalpure,Akshata S.Gavde,Pranjali Choudary,"Voice assistant using Python",International Journal of Creative Research Thoughts,Volume 10,issue 6,June 2022,page no. 801-806.

[2] Jash Vora, Deepak Yadav, Ronak Jain Jaya Gupta,"JARVIS: A PC Voice Assistant",International Journal of Advance Study and Research Work,Volume 4,Issue 7,July 2021,page no.386-393.

[3] Dr. Jaydeep Patil, Atharva Shewale, Ekta Bhushan, Alister Fernandes, Rucha Khartadkar,"A Voice Based Assistant Using Google Dialogflow and Machine Learning",International Journal of Scientific Research in Science and Technology,Volume 8, Issue 3, May-June2021,Page No. 06-17.

[4] Abeed Sayyed, AshpakShaikh, AshishSancheti, Swikar Sangamnere, Prof. Jayant H Bhangal,"Desktop assistant using python",International Journal of Advanced Research in Science,Volume 6, Issue 2, June 2021,page no.1327-1335.

[5] Anjali Fapal, Trupti Kanade, Bharati Janrao, Mrunalini Kamble,Megha Raule,"personal voice assistant windows using python",International Research Journal of Modernization in Engineering REFERENCES Technology and Science ,volume 3,issue 7,july 2021,page no.485-491.

[6] V. Geetha, C.K.Gomathy, Kottamasu Manasa Sri Vardhan, Nukala Pavan Kumar,"The Voice Enabled Personal Assistant for Pc using Python",International Journal of Engineering and Advanced Technology (IJEAT) ISSN: 2249-8958, Volume 10 Issue 4, April 2021,page no.162-165.

[7] Subhash Mani Kaushal, Megha Mishra,"A Review on Voice Assistance using Python",International Research Journal of Engineering and Technology (IRJET),Volume 07 Issue 09,Sep 2020,Page no.3862-3864.

[8] Ankit Pandey, Vaibhav Vashist, Prateek Tiwari, Sunil Sikka,Priyanka Makkar,"Smart Voice Based Virtual Personal Assistants with Artificial Intelligence",Artificial Computational Intelligence,Volume 1,Issue 3,June 2020,page no.385-397.

[9] ALI BOU NASSIF , ISMAIL SHAHIN , IMTINAN ATTILI,MOHAMMAD AZZEH2, AND KHALED SHAALAN,"Speech Recognition Using Deep Neural Networks: A Systematic Review",February 1 2019,VOLUME 7,page no.19143-19165.

[10] Nasrin Aktar , Israt Jahan , Bijoya Lala "Voice Recognition based Intelligent Wheelchair and GPS Tracking System ",2019 International Conference on Electrical, Computer and Communication Engineering (ECCE), 7-9 February, 2019 .

[11] Dr. Kshama V. Kulhalli,Dr.Kotrappa Sirbi,Mr. Abhijit J. Patankar,"Personal Assistant with Voice Recognition Intelligence",Volume 10, Number 1,Issue 2017,pg no.416-419.