INTERPRETER OF VOICE AND ACTION EXECUTER

Samay Chakra[1], Sanaboina Vishnu Vardhan[2], Sarvasetty Sai Satya Shreyas[3], Ayesha Taranum[4]

[1] *Student*, Computer Science & Engineering, Presidency University, *Karnataka, India*
[2] *Student*, Computer Science & Engineering, Presidency University, *Karnataka, India*
[3] *Student*, Computer Science & Engineering, Presidency University, *Karnataka, India*
[4] *Assistant Professor*, Computer Science & Engineering, Presidency University, *Karnataka, India*

ABSTRACT

As Python is an emerging language, writing a program for Virtual Assistant using Python becomes easier. Instructions for the assistant can be handled as per the user requirement. Text is converted to speech using Speech recognition. This feature is commonly used in voice assistants such as Alexa, Siri, etc. Python consists of an API known as SpeechRecognition which facilitates the conversion of speech to text. Making a personal voice assistant is a really interesting task. Sending emails became easier as you didn't have to type a word, you didn't have to open a browser to search on google, and performing many other daily tasks like playing music and opening your favourite IDE with the help of a single voice command. Advancement in technologies is such that any task can be performed with similar effectiveness or it is safe to say that more effectively than us.

Keyword: - *Desktop Assistant, Voice Assistant, Python, and Text to speech*

1. INTRODUCTION

What is a Voice Assistant? A voice collaborator is a menial helper that perceives voice and uses calculations that interaction language and incorporates client voice to pay attention to explicit orders and return wanted yield that is significant or carry out unambiguous roles as mentioned by the client. In view of explicit orders, in some cases alluded to as aims, verbally expressed by the client, voice partners return applicable data by tuning in for explicit catchphrases and sifting through the encompassing clamor. This venture depends on windows programming advancement and give individual collaborator utilizing voice acknowledgment. This program incorporates capacities, for example, sending instant message, email administration, WhatsApp administration, Instagram administration. Screen capture administration, YouTube administration, area administrations, playing music, checking climate, Google web crawler, Wikipedia web search tool and so on. As voice collaborator coordinates the majority of the framework administrations which are utilized consistently, it could help individuals having handicaps for manual tasks and help in getting a more advantageous life. Fundamental goal of building individual voice collaborator programming is to utilize the semantic information sources that are accessible on the web, client created content and to give information from information data sets. The primary motivation behind a menial helper is to respond to questions that clients might have. This might be finished in a business climate, on the business site, with a visit interface. While utilizing the portable, the savvy remote helper is accessible as a call-button worked administration where a voice asks the client "How might I at any point help you?" and afterward it answers verbal information. You can utilize the voice aide to get any of your undertakings required simply by programming it to do as such. The primary benefits of voice search is its speed. As a matter of fact, voice search can be multiple times quicker than a composed pursuit as we can expound on 40 words each moment, we are equipped for talking around

150 words inside a similar time span. In this regard, the capacity of individual partners to precisely perceive expressed words is an essential for them to be taken on by purchasers.

2. LITERATURE SURVEY

Abhay Dekate (2016) et al. introduced that thing that were once viewed as incomprehensible should be possible now at the same time, to have the option to accomplish and achieve these considerations a stage is required which will actually want to play out our errands effortlessly. The solicitation as sound is caught by means of amplifier and the solicitation is handled so the gadget can answer the client with the utilization of the speaker module.[1]

Rutuja V. Kukade (2018) et al. suggested that visually impaired individuals had a great deal of correspondence boundaries, and they face different difficulties because of this. In this paper, execution of an individual voice partner has been examined which will actually want to take the human information by means of voice orders to execute task.[2]

Deny Nancy (2019) et al. introduced the chance of tackling errands that were once viewed as unimaginable in the Modern Era effortlessly and comfort.[3]

Tushar Gharge introduced the issues experienced by client while fostering a PC program. Fostering a PC program is definitely not a simple errand as it needs equipment assets which client need to deal with. There might be plausible of wounds to the fingers of the client while composing the code persistently. To stay away from issues thusly, a product is to be planned in which utilizations voice to get the undertakings done.[4]

Veton Këpuska (2018) recommended that people and machines should understand regular language and that is one of the objectives of Artificial knowledge (AI). As of late, the intuitive conversational frameworks are the quickest developing region in AI.[5]

Bassam A Raja N, composed an explanation about discourse which is generally huge. In this, the correspondence among human and machine was finished by utilizing a simple sign which is changed over by discourse sign to computerized wave.[6]

The verbal enunciation is transformed from a person into a substitute technique for information (e.g., text) utilizing a recognizer.[7]

J. B. Allen portrayed about the language that is the greatest amount of huge method for correspondence and discourse is its significant connection point. The point of interaction for human to machine, discourse signal was changed over into simple and computerized wave shape as a machine understood.[8]

Mugdha Bapat, Pushpak Bhattacharyya, portrayed a morphological analyser for the vast majority of the NLP sales of Indian Languages.[9]

Sean R Eddy worked on Hidden Markov models which are a typical factual planning approach for 'straight' issues like groupings or time series and have been widely used in discourse distinguishing proof solicitations for as much as twenty years.[10]

An Intelligent Voice Assistant framework was created as an android application that showed the utilization of normal language handling (NLP) which assists with sending messages and even utilize the in-constructed versatile application by the use of voice orders. This framework was adjusted to involve the mailing and schedule capacities wherein the client had the option to mail and make their occasion utilizing voice order. Home Automation framework in view of Internet of Things worked acceptably by the association of basic apparatuses to it. The

application apparatuses were effectively controlled somewhat through web. The planned framework doesn't just screen the sensor information like temperature, gas, light, movement sensors yet additionally it incites a cycle as per the necessity. For example, how about we consider turning on the lights when it gets dull. Sensor boundaries are likewise put away in the cloud as soon a possible. This will assist client with dissecting the state of different boundaries in the home whenever anyplace. Everybody should be known all about Cortana, Siri, Watson or Google Now or with any of the incalculable remote helper. These genuine menial helpers aren't quite as shrewd as Ironman's Jarvis; however, their planned capacity is practically something very similar, voice-enacted processing that is controlled by man-made consciousness. Pose an inquiry and find a solution. Provide an order and come by results.

Amazon's Alexa is an associate that figures out how to separate itself. Dissimilar to versatile based remote helpers like Siri, Alexa is unified inside committed, in-home Amazon gadgets, most remarkably the Amazon Echo, a consistently on, continuously listening Internet associated speaker. A few highlights of Alexa incorporate streaming music, perusing the titles and responding to your inquiries.

3. EXISTING SYSTEM

Individual collaborator on work area is fabricated utilizing AI advances, end-client is helped by the product with everyday exercises like general human discussion, looking through questions in web search tools, looking for recordings in YouTube, documents in nearby drive, recovering pictures, playing melodies in work area by utilizing imported modules like pyttsx3 is a disconnected module that is utilized for text to discourse transformation in python, Date and time usefulness is upheld by utilizing the DateTime module. The client articulations are breaking down with the assistance of Artificial knowledge to give an ideal arrangement. In spite of having different advantages given by discourse acknowledgment, the framework is additionally having constraints. By embroiling the advancement of discourse acknowledgment these limits are additionally acquired by the applications. Design acknowledgment is utilized by the generally existing Voice Assistants which comes up short on setting, the precision, and misinterpretations, time, expenses, efficiency and client emphasizes. They work just on web-based mode. Information is put away in the data set servers which prompts expansion in Time and Space Complexity. Putting away information in cloud prompts security issues. Foundation commotion obstruction is additionally one more issue with discourse acknowledgment programming.

4. PROPOSED SYSTEM

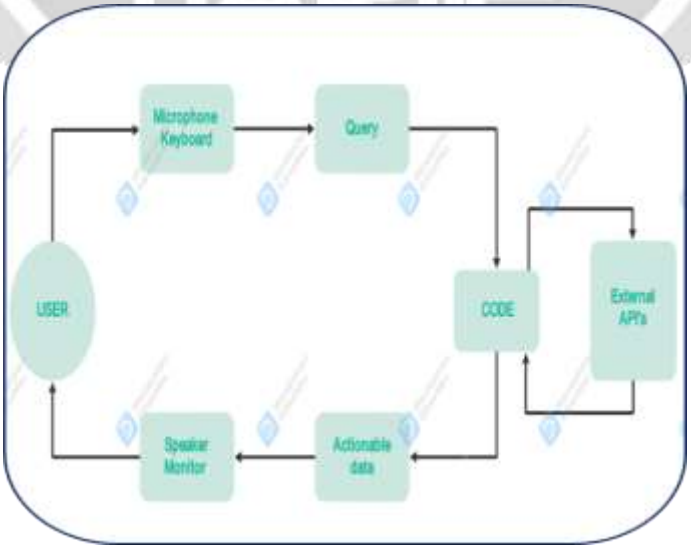
- In the event that framework neglects to assemble data from the client input, it will continue to ask in the future to rehash till it distinguishes the client input.
- Both male and female voices can be utilized by the framework as per client necessities.
- Will keep password.py record which will ask the client for secret key prior to permitting it to get to the primary capacity of the document.
- In PC, voice partner will actually want to finish undertakings like playing music for us, opening sites, showing area, opening framework applications, tackle assignments like sending messages and messages.
- Use hotword to awaken the voice right hand.
- Voice acknowledgment can't be amazing because of the highlight however it very well may be made more effective by utilizing the local complement (en-in)
- As you will program the right hand without help from anyone else, the information needn't bother with to be put away and consequently there won't be a break in protection.
- Hotword is made to begin the voice collaborator simply by expressing it to awaken.
- Secret phrase safeguarded in order to let just the approved individuals to get to.

5. TECHNOLOGIES USED

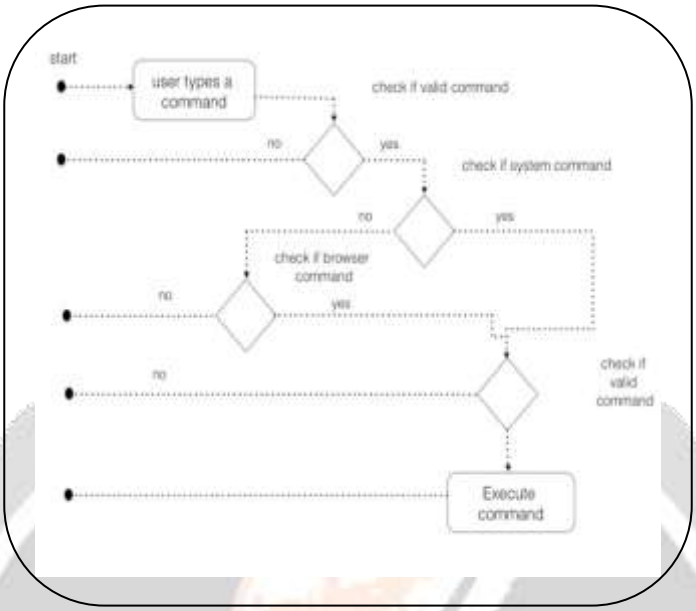
Python is the language used to program this undertaking, as it is flexible, and it gives a wide scope of library. Programming the Voice Assistant is finished utilizing Microsoft Visual Studio Code (IDE) which upholds Python. To perceive the client input, SpeechRecognition module is utilized which comprises of capacities that are in-constructed. We initially characterize a capacity to change the text over completely to discourse. Pyttsx3 library is utilized for something very similar. Then the library example is instated into a variable. We use say () strategy and pass the text as a contention to that, for which it gives a voice answer as a result. To perceive the voice order given by the client, another capacity has been characterized. In that capacity, we characterize magnifying lens source and inside its degree, we utilize particular capacities and store the result in a variable. We have many administrations to use for this entire reason, similar to Google Speech Recognition motor, Microsoft Bing Voice. For this venture, we pick Microsoft's Speech Recognition Engine, that will change over the voice order into a computerized text design. That text is then passed as a contribution to the Assistant, and afterward the associate beginnings looking for the watchword. Assuming the information order has a word that coordinates with the individual word, the particular capacity will be called, and it will play out the activity as needs be, such as telling the date, or time, or taking a screen capture, or saving a short note, and some more. The primary benefit of this Personal Virtual Assistant, is that it saves a great deal of time, and it can deal with questions even from individuals, of various voices. There is no prerequisite as to such an extent that one needs to provide any definite determined order to set off a specific activity. Client has the adaptability to provide order for client, in regular language.

6. SYSTEM DESIGN

- Contribution from the client is taken as voice.
- Discourse is then changed over into text to be handled by the collaborator.
- Results are currently acquired by handling the changed over texts.
- The contribution from the client contains catchphrase which decides the undertaking that must be performed. On the off chance that the catchphrase doesn't match any of the questions in the code, then the partner requests that the client talk once more.
- Results got by the framework in text design is then switched over completely to discourse configuration to create the last result.



ARCHITECTURE DIAGRAM



FLOW CHART

6.1 SYSTEM IMPLEMENTATION

Allow us to isolate the Software into 6 fundamental stages prior to discussing execution.

The main stage is "Actuation", so IVAE gets enacted once it hears the watchword "WAKE UP".

The subsequent stage is to enter the secret key to get to IVAE.

The third stage is "Discourse to-message", IVAE will begin recording what the client is talking about then convert it to message.

The fourth stage is "Aim acknowledgment", it will attempt to sort out the transition to the following stage. If not, it will request that the client say that in the future and return to the discourse to-message.

The fifth stage is "handling and execution", which will attempt to make the legitimate move, if any, in light of the aim of what has been said. Any other way, it will likewise let the client know that it couldn't make a legitimate move and solicitation the client to rehash the order, and afterward, return to the discourse to-message part.

The 6th stage is to change over "Text-to-discourse". At this stage, IVAE will either recognize or give a reaction to what has been mentioned.

7. RESULTS AND DISCUSSION

Remote helper consumes insignificant opportunity to deliver the ideal result. Menial helper is a program which comprehends composed/verbal orders and executes the errand doled out by client. Natural language Processing (NLP) is utilized to coordinate client voice or text input with executable orders. Utilizing the voice colleague, you are equipped for running your framework like PC or PCs on your own order. It saves time as it is a quick interaction. Menial helper works at set times for you, so it is generally accessible to you and it is competent to adjust to changing requirements rapidly.

8. CONCLUSION

This framework is planned in such a strategy wherein the client can oblige to it easily. Our proposed framework VPA, an individual voice collaborator can be executed utilizing SpeechRecognition module, subsequently making the framework safer and powerful. VPA contributes the voice control application that gives improvements to all applications running on a framework by incorporating orders set from onscreen setting. VPA can help enormous number of clients with general eyes free and sans hands voice control of their framework. Discourse acknowledgment is a key innovation which will give a better approach for collaboration between human, machine and instruments. While interfacing with a non-visual screen, voice orders is more profitable over multi-contact as there is no need of focuses to find it. The sending of email, and perusing of information will likewise be feasible for the visually impaired individuals. This can-do assortment of errands like let you know the time, open application, coordinated records, can give updates of matches, play game, let you know the area, make a few quips, open hackathon, do estimation, refreshes about the stock and the vast undertakings for the client. Subsequently, it makes one's life agreeable and simultaneously somewhat open through voice orders.

9. FUTURE SCOPE

Versatile App Integration: The most sweltering pattern right currently is to incorporate voice-tech into portable, and a similar will stay as voice is a characteristic UI (NUI). Usefulness is expanded by voice fueled applications, and it saves clients from convoluted application route. On the off chance that an individual doesn't know precisely exact thing or where to find, voice enacted applications make it simple to explore the inquiry. At this stage it seems like voice coordination might appear to be good to have however it will before long turn into a compulsory necessity.

Voice Cloning: It makes the PC created voice so indistinct from the client's voice, which gives the discourse more profound by utilizing AI (ML) tech and GPU power improvement. You simply need to utilize a discourse that is pre-recorded and afterward a voice change innovation will change your voice into another.

Add neural network which will make the framework more astute and give it its own practical mind, where-in it can learn while bantering with client.

10. REFERENCES

[1] Abhay Dekate, Chaitanya Kulkarni, Rohan Killedar, “Study of Voice Controlled Personal Assistant Device”, International Journal of Computer Trends and Technology (IJCTT) – Volume 42 Number 1 – December 2016.

[2] Rutuja V. Kukade, Ruchita G. Fengse, Kiran D. Rodge, Siddhi P. Ransing, Vina M. Lomte, “Virtual Personal Assistant for the Blind”, International Journal of Computer Science and Technology (JCST), Volume 9, Issue 4, October - December 2018.

[3] Deny Nancy, Sumithra Praveen, Anushria Sai, M. Ganga, R.S. Abisree, “Voice Assistant Application for a college website”, International Journal of Recent Technology and Engineering (IJRTE) ISSN: 2277-3878, Volume-7, Issue-6S5, April 2019.

[4] Tushar Gharge, Chintan Chitroda, Nishit Bhagat, Kathapriya Giri, “AI-Smart Assistant”, International Research Journal of Engineering and Technology (IRJET), Volume: 06 Issue: 01, January 2019.

[5] Veton Këpuska, "Next-Generation of Virtual Personal Assistants (Microsoft Cortana, Apple Siri, Amazon Alexa and Google

[6] M. Bapat, H. Gune, and P. Bhattacharyya, "A paradigm-based finite state morphological analyzer for marathi," in Proceedings of the 1st Workshop on South and Southeast Asian Natural Language Processing (WSSANLP), pp. 26–34, 2010.

[7] Sarikaya, R.: The technology behind personal digital assistants. IEEE Signal Process. Mag. 34, 67–81 (2017).

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

[9] M. Bapat, H. Gune, and P. Bhattacharyya, "A paradigm-based finite state morphological analyzer for marathi," in Proceedings of the 1st Workshop on South and Southeast Asian Natural Language Processing (WSSANLP), pp. 26–34, 2010.

[10] S. R. Eddy, "Hidden Markov models," Current opinion in structural biology, vol. 6, no. 3 3, pp. 361–365, 1996.