INTERNATIONAL JOURNAL
FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

www.ijraset.com

Call: ⓒ08813907089 | E-mail ID: ijraset@gmail.com

Virtual Voice Assistant

Ch. V. Tejaswi[1], B. Sravya[2], D. Ratnam[3], A. Jeevanjali[4], T. SriLakshmi[5]

[1, 2, 3, 4]UG Students, [5]Assistant Professor, Computerr Science and Engineering, PVP Siddhartha Institute of Technology

Abstract: This is desktop application which can assist people with basic tasks using natural language. Virtual Voice Assistants can go online and search for an answer to a user's question. Actions can be triggered using text or voice. Voice is the key. A virtual voice assistant is a personal assistant which uses natural language processing (NLP) , voice recognition and speech synthesis to provide a service through a particular application. Natural Language Processing in short is called as NLP. It is basically a branch of artificial intelligence which mainly deals with the interaction between personal computers and human beings using the natural language. The main objective of NLP is to read, convert, understand, and make use of the human languages in a manner that is valuable. Voice recognition is a hardware device or computer software program with the potential to decode the voice of human beings.

Voice recognition is usually used to operate a gadget, execute commands, or write without making use of any mouse, keyboard, or press any buttons. Artificial production of human speech is called as Speech Synthesis. A system used for this purpose is called a speech computer or speech synthesizer and can be implemented in many products of software's and hardware's.

Keywords: Voice, User , Command, Assistant , Nova

I. INTRODUCTION

For most of us, the eventual comfort would be an assistant who always listens for your call, foresee your every need, and takes action when necessary. That luxury is now available for everyone, thanks to artificial intelligence assistants, aka voice assistants.

Virtual Voice Assistant is an application that can understand voice commands and complete tasks for a user. The medium of communication we are using is English Language. We named our Virtual Voice Assistant as NOVA.

Voice assistants come in a bit small packages and they can perform a variety of actions after hearing a stir word or command. They can answer questions, play videos, sends whatsapp messages, etc. NOVA can help someone with basic tasks. They often understand natural language and can help with things like playing videos, telling a joke etc.

NOVA have access to a large amount of information on a device or online, which enables them to perform simple tasks. Humans often find it difficult and intimidating to do their own tasks manually such as opening web browser, searching content online, playing youtube videos, typing.

In order to make it more convenient we do it by voice commands rather than manual typing which can save time. The application will start with voice mode as its primary mode for the voice assistant so that the user can start interacting with the application immediately.

An easy way to comply with IJRASET paper formatting requirements is to use this document as a template and simply type your text into it.

II. LITERATURE SURVEY

Voice recognition technology was found long ago Apple's Siri in 2011. At the Seattle World's Fair in 1962, Shoebox is a tool presented by IBM. It was the size of a shoebox and which can perform mathematical functions and Acknowledge 16 spoken words as well as digits 0-9.

Scientists at Carnegie Mellon University in Pittsburgh, Pennsylvania — with the substantial support of the us Department of defence and its Defence Advanced Research Projects Agency (DARPA) — created Harpy ,In the 1970s. It can nearly acknowledge 1,011 words, which is about the vocabulary of a three-year-old.

Once organizations came up with contrivance that could acknowledge word sequences, companies began to build applications for this technology. The Julie doll from the Worlds of Wonder toy company came out and could recognize a child's voice and respond to it in 1987.

In 1990s, companies like Apple, IBM, and others also created items that used voice recognition technology for best invention. Apple began building speech recognition features in 1993.

Macintosh computers with Plain Talk. Dragon came out with Dragon NaturallySpeaking, which was the first continuous dictation product, In April 1997. It can understand more than 100 words per minute and turn it into text. One of the earliest adopters of voice recognition technology was Medical dictation devices.

III.PROPOSED WORK

A. Training the wake Words

Wake words depend on a special word which is always listening for a particular word or phrase so that a phone, smart speaker, or something else can begin communicating with a server to do its job. Stir words need to be long enough to be clear, it is easy for a human to speak, and simple for a machine to acknowledge. This is why you cannot change your stir word to anything you want it to be. Voice assistants don't really "realize" what you're saying — they just listen for their wake word and then begin interacting with a server to finish a job..

B. Running the Application

After training the wake words to the application, now it can understand those wake words and NOVA can give response to those words which was requested by the end user in order to perform the tasks. By running the application NOVA starts its communicating with the end users. NOVA is the best interface to the users. If the wake words requested by the user is fluent and correct then NOVA will understands it and give response according to the requested command. Based on the speed of the internet a NOVA can give fast response to the end user. If the requested commands are not wake words or not pronounced correctly the NOVA will produce the error.
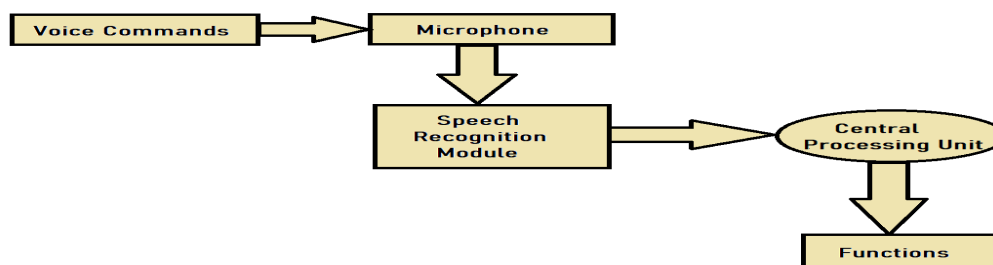
C. Basic Work Flow

Fig. 1 Work flow

Different functions and services included in our application are :

1) *Text-to-Speech:* Text-to-speech (TTS) is basically a type of assistive technology that study digital text aloud. It's sometimes called "read aloud" technology. Just with a button click , TTS takes words on a system or any other digital device and converts them into speech. TTS can enable the reading of computer that exhibits the information for the visually challenged person, or may simply be used to augment the reading of a text message.
2) *Playing Youtube Videos:* We can play different Youtube videos simply by giving voice commands.
3) *Searching Content Online:* We can search for content online easily by specifying voice commands.
4) *Speech-to-text:* Speech-to-text is a module which effectively takes audio content and interprets it into written words in a word processor or other exhibit destination. This type of speech recognition software is exceptionally valuable to anyone who needs to generate a pile of written content without a lot of manual typing. This is also very useful for people with dysfunctions that make it difficult for them to use a keyboard.
5) *Remembering Things:* Our Voice Assistant is also capable of remembering things for the user whenever the user tells to remember.
6) *Striking Jokes:* Joke is something which can provide fun, cause amusement or laughter. Our voice assistant can also make the users laugh by cracking jokes.
7) *Date Time:* User can also get the updates about date and time using our voice assistant.
8) *Delivering Whatsapp messages:* We can also send whatsapp message at certain time. It is one of the easiest way for scheduling WhatsApp messages.

IV. RESULTS AND EVALUATION

After the application is built , now we need to test it against various modules. Testing is done and both positive and negative testcases are evaluated. The principles followed for testing are

All tests should be traceable to user requirements

1) The positive flow of the task should be taken in consideration.
2) Testing is a process which should start "from the beginning" and progress towards testing "to the end"
3) Exhaustive testing is not possible

A. *Positive test Cases*

The positive flow of the functionality must be considered

Valid inputs must be used for testing

Must have the positive view to verify whether the requirements are justified.
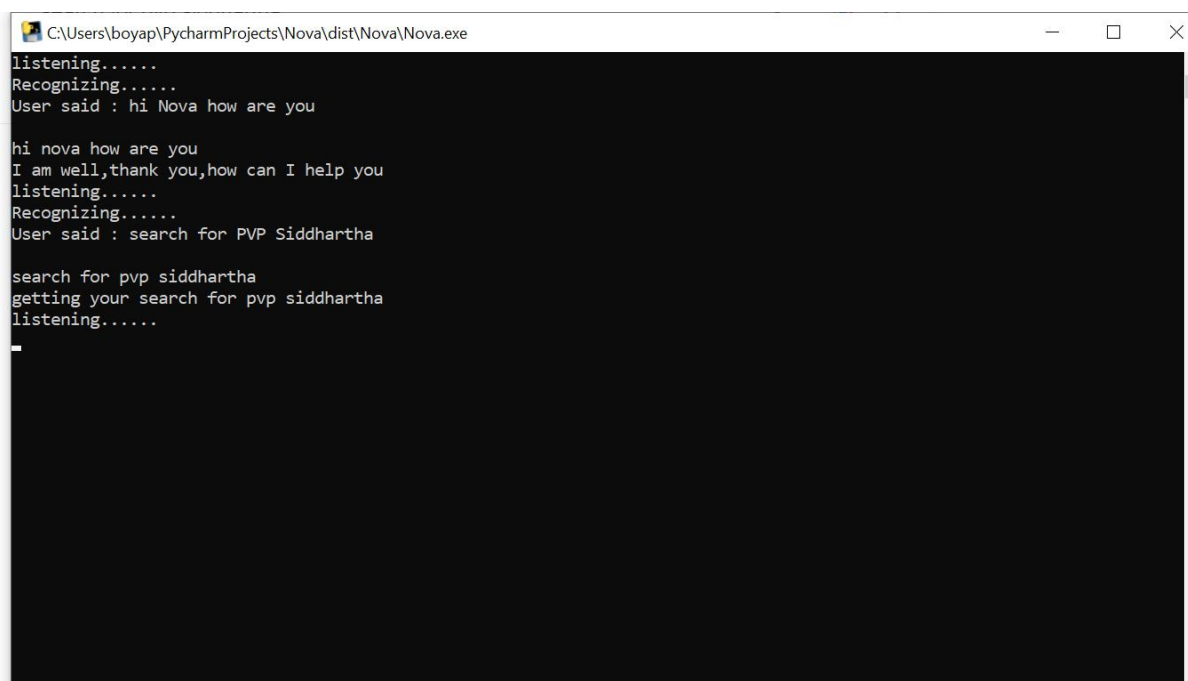
TABLE I
POSITIVE TEST CASES

Test cases	Description	Input	Expected Output	Actual Output	Result
Pyttsx3	Converts the text entered, into audio	Text	Text to speech converted.	Text to speech converted	Success
Speech Recognition	Recognises the speech spoken by the end user	Voice	Speech recognised	Speech recognised	Success
Web browser	Used to open web browser	Input search query	Opens web browser	Opens web browser	Success
Date Time	Displays the date and time		Date and time displayed	Date and time displayed	Success
Playing youtube videos	Plays youtube videos	input video name	Youtube Video Played	Youtube Video Played	Success
Remember--ing things	Remembering things for the user	Input query	remembered	remembered	Success
pywhatkit	Used to send Whatsapp messages	receiver phone number	Message sent	Message sent	Success

TABLE III

NEGATIVE TESTCASES

Test cases	Description	Input	Expected Output	Actual Output	Result
Trained Keywords	NOVA can recognise only the trained words	Voice	Response for the requested commands	Not Recognised	Failure
Language	NOVA only understands the American English	Voice	Response for the requested commands	Not Recognised	Failure
Whatsapp messages	NOVA can send message through voice commands if the phone number is wrong	Voice	Send message	Not send	Failure

Here invalid inputs must be used for testing.

Screenshots illustrating the interface of our application :
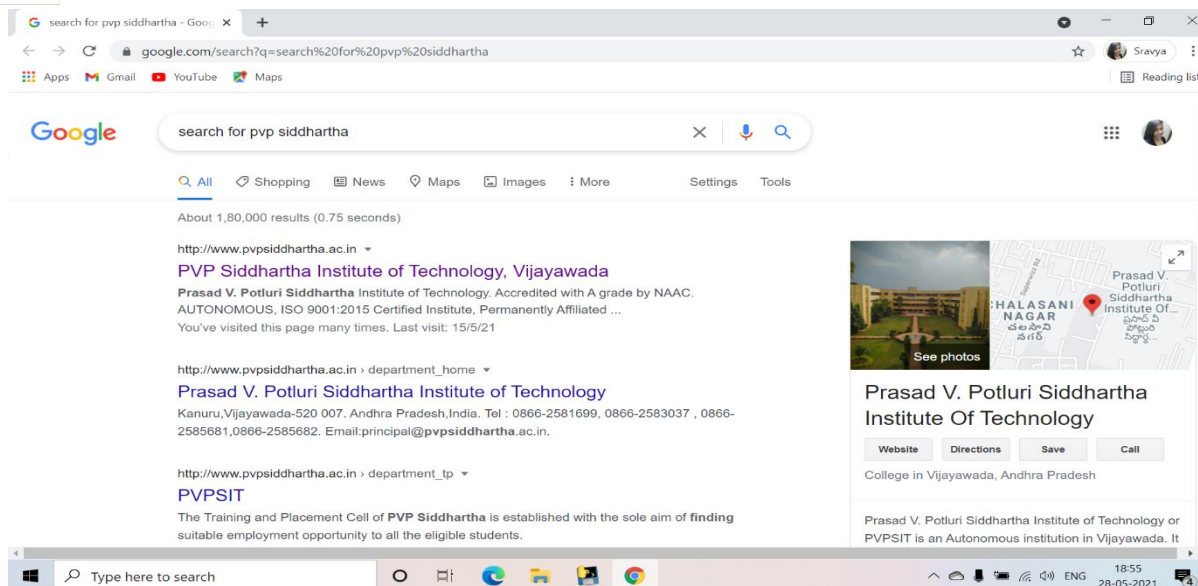

Fig. 1 Search query

Fig. 3 Open browser

V. CONCLUSIONS

Virtual assistants are evolving quickly. Companies are enabling them to provide more abilities like speech recognition and natural language processing advances. It enables them to understand and perform requests.

The world of virtual voice assistants was born to make life easier for humans. Our job can be solved in a matter of seconds just by using our voice. With more accurate virtual assistants, we will see more Voice Market strategies.

Voice assistants come in packages and can perform a variety of actions after hearing a wake word or command. They can answer questions, play music, play videos and tell jokes etc.

Our goal is to create a virtual assistant that provides a better interface to the user which can be very flexible and useful to any user .Our aim is to develop and propose a system used for the end users communicating through voice commands.

VI. FURTHER SCOPE

NOVA can be extended in future by importing some more modules for performing different tasks. We can also train the NOVA to learn some more skills by its own. By taking the NOVA in the real world we can make our lives more effective and easier. By adding the EmoVoice module in the NOVA we can catch the emotion from the voice commands of the end user, based on that NOVA will respond to the user. By using the imageEnhance module we can also adjust the brightness of the desktop. By using PyAudio we can control the volume of our desktop through voice commands itself.

REFERENCES

[1] Arya, A Progeniture Intelligent Assistant For Superior User Experience, April 2020, https://eprajournals.com/jpanel/upload/133am_80.EPRA%20Journals%20-4335.pdf
[2] Subhash Mani Kaushal, Megha Mishra, A Review on Voice Assistance using Python, September 2020 https://www.irjet.net/archives/V7/i9/IRJET-V7I9678.pdf
[3] Analytics vidhya, Build Your Own Desktop Voice Assistant in Python, November 3 2020 https://www.analyticsvidhya.com/blog/2020/11/build-your-own-desktop-voice-assistant-in-python/
[4] Diana Ramos, Voice Assistants: How Artificial Intelligence Assistants Are Changing Our Lives Every Day, April 16, 2018 https://www.smartsheet.com/voice-assistants-artificial-intelligence
[5] Investigating the Accessibility of Voice Assistants With Impaired Users: Mixed Methods Study , November 9,2020
[6] https://www.jmir.org/2020/9/e18431/Olmstead K. Nearly half of Americans use digital voice assistants, mostly on their smartphones.Washington: Pew Internet and American Life Project , December 12 2017,https://www.pewresearch.org/fact-tank/2017/12/12/nearly-half-of-americans-use-digital-voice-assistants-mostly-on-their-smartphones/
[7] Forbes, Voice Assistants : This is what the future of technology looks like , November 1,2017,https://www.forbes.com/sites/herbertrsim/2017/11/01/voice-assistants-this-is-what-the-future-of-technology-looks-like/?sh=419d3f29523a

INTERNATIONAL JOURNAL
FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 ⊘ (24*7 Support on Whatsapp)