



Desktop Voice Assistant Using Natural Language Processing (NLP)

Lalit Kumar

Department of Information Technology, Maharaja Agrasen Institute of Technology, Delhi, India

To Cite this Article

Lalit Kumar, "Desktop Voice Assistant Using Natural Language Processing (NLP)", *International Journal for Modern Trends in Science and Technology*, 6(12): 332-335, 2020.

Article Info

Received on 12-November-2020, Revised on 05-December-2020, Accepted on 11-December-2020, Published on 15-December-2020.

ABSTRACT

Voice assistants are the great innovation in the field of AI that can change the way of living of the people in a different manner. the voice assistant was first introduced on smartphones and after the popularity it got. It was widely accepted by all. Initially, the voice assistant was mostly being used in smartphones and laptops but now it is also coming as home automation and smart speakers. Many devices are becoming smarter in their own way to interact with human in an easy language. The Desktop based voice assistant are the programs that can recognize human voices and can respond via integrated voice system. This paper will define the working of a voice assistants, their main problems and limitations. In this paper it is described that the method of creating a voice assistant without using cloud services, which will allow the expansion of such devices in the future.

KEYWORDS: *Voice Assistant, Natural Language Processing, Speech Recognition, Python.*

I. INTRODUCTION

The development of artificial intelligence is growing day by day that are able to recognize human voice as a command, analyse it and provide services to humans. The voice assistant is gaining lots of popularity today. It is the era where no longer a human learns how to communicate with machines?, but a machine learns to communicate with a human, exploring his actions, asking about his hobbies, habits and trying become a human's personal assistant.

A voice assistant is a software program that can perform tasks or provide some kind services for an individual based on the human verbal commands i.e. by using human voice commands and the voice assistant will respond via synthesized voice. The Users can ask their assistants' questions, control home automation devices, and media playback via voice, and manage other basic tasks such as email,

to-do lists, open or close any application, send messages to anyone on WhatsApp etc. with verbal commands only. Some other types of Voice Assistant are:

- Intelligent Personal Assistant
- Automated Personal Assistant
- Virtual Digital Assistants
- Chatbot

There is great example of the voice assistant on Desktop, Cortana (Voice Assistant) which is an intelligent personal assistant, human language interface, and voice recognition software for Windows PC. Cortana is a multi-functional AI Voice Assistant that allows user to interact with their computer using voice commands and it is available in most of the languages of the world. Cortana also supports the accurate conversion of speech to text in over 100 different languages of the world.

Today most of us just want a program like Voice

Assistant that listen to us, perform some tasks like attend call for us, open a particular application, anticipate you every need and take action when necessary. Well that technology is now available and all thanks to artificial intelligence assistants, aka voice assistants. Voice assistants come in around small packages and can perform a variety of actions after hearing person's command. They can turn on lights, answer questions, play music, place online orders and do all kinds of AI-based stuff.

Based on voice assistant concept a well-known application is also developed by the Google that is "Google Voice Assistant" which is used for in Android Phones. Also, we have similar applications like Bixby (personnel assistant which is only available in Samsung devices), Cortana, and Google Voice Assistance.

II. BACKGROUND

A. History of Voice Assistants

If we look into the history of Voice Assistant, the first voice activated device was 'Radio rex' in 1911 and in 1964 IBM displayed another unique apparatus called shoebox. it was the actual size of the shoebox and it could perform scientific functions, can perceive 16 words. The Foundation of smart Voice Assistant was laid by IBM Simon as we know today. During 1970s, there was another research project going on at Carnegie Mellon University in Pittsburgh, Pennsylvania - with the help of U.S. Department of Defense Advanced Research Projects Agency (DARPA) and they introduced Harpy which could understand 1000 words, which is approximately the vocabulary of a three-year-old child. In 90s Some big organizations like Apple, Google, IBM started working on voice acknowledgment. in 1993, Macintosh began to build speech recognition with its Macintosh PCs with PlainTalk.

B. Future Application

In coming future there will be so many technologies in which Voice Assistant will be used. Some of the examples are in companies, voice assistants can be used to automate repetitive tasks, for example Amazon's Alexa will open scheduled video conference and book meeting rooms. In smartphones your voice assistant will be able to talk on your behalf while you are away from your phone or doing some work. Even it will send messages also.

C. Aim of this project

The aim of this project is that we have created a program or intelligent personal assistant which will perform some tasks like sending messages to WhatsApp, opening applications, giving information about the weather etc. in response to the voice commands which will be provided by the users.

III. RELATED WORK

Every company who are perfect in intelligent voice assistant has applied his own methods and approaches for development, which in turn affects the final product. Voice assistant using Natural Language Processing provides a variety of services using artificial intelligence systems equipped through the user's voice commands. According to the artificial intelligence system, the system is divided into two main parts the one is platform-management type and another is open API type AI. Platform-management type AI can able to do detailed operation, but it is not possible to use the services on other platforms as it is dependent on one platform. And open API type AI, it's possible to use services on other platforms. When we talk about voice assistants the first thing comes in our mind is some of the great application in this field like Google Assistant, Cortana, Siri, amazon's Alexa etc. which are mainly controlled by users voice commands. Once user sends a voice command to these devices, it processes into text, converts the data processed into speech after returning it and passes it to the output device. Some other main technologies related to this field are voice activation, automatic speech recognition, Text-To-Speech, voice translator, named entity recognition.

IV. PURPOSE PLAN OF WORK

The proposed plan started by providing voice input to the voice assistant by the user through microphone which later processed and analyzed by voice assistant. the voice input can be anything like getting any information, operation on computer's internal files etc. This is the study based on reading above mentioned literature and testing their examples. Speech recognition has been used to convert the voice input into text. Then this text is then passed through the central processor which analyses the purpose of the command and calls the required script for execution.

Even though there are hundreds of inputs have been provided to the program, but the complexities

don't stop here. Other factors like Background noises, same kind of voices can also play huge role in not understanding the voice inputs by the voice assistant. the main reason behind is that it does not have the ability to distinguish the unwanted sounds it hears. For that the developer have to train the model and add it into the voice assistant program and for model training they collect different data of these ambient sounds and tell the program to filter them out. Another factor which may come in action is the shift in pitch of human voice to make it easy to recognize by voice assistant.

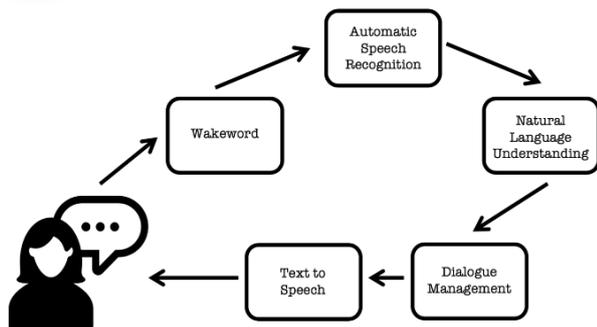


Fig. 1 Shows the Adding of wake word by which Voice Assistant will Activated.

V. METHODOLOGY

Voice recognition is a relatively narrow field. This means given enough samples, you can create a model that can recognize and transcribe voice commands under different circumstances and with different background noises and accents. Some features which are used:

A. Speech Recognition

The system is using Google's online speech recognition system for converting speech input to text. Through this the users can speak and obtain the text in exchange of voice input from the special corpora organized on computer network server at the information center from the microphone which is temporarily stored in the system and then sent to google cloud for speech recognition. the same text is then received and sent to the voice assistant program.

B. Python Backend

The whole program is written on python Backend. The python Backend work on getting the output in exchange of voice input provided by the users through speech recognition module and then identifies whether the command given is Context Extraction, API Call and System call. The response is then sent back to give required output.

C. API Calls

API stands for Application Programming Interface. It is software intermediary and the work of the API is to allow two applications to talk each other. In other words, API is the message passer that delivers the users request to the provider and then send the response back to users.

D. Context Extraction

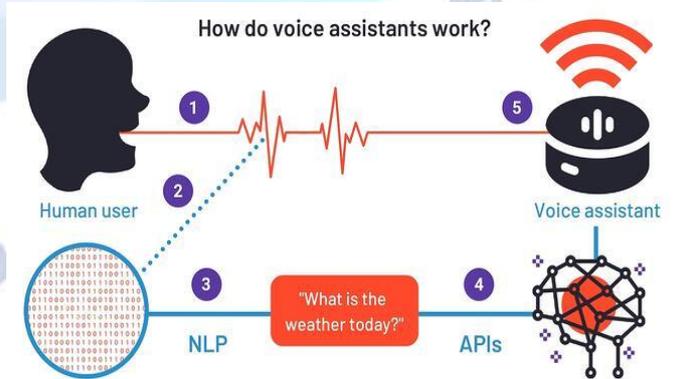
Context Extraction is use to automatically extract structured information from unstructured or semi-structured machine-readable documents. This task uses natural language processing (NLP) for processing human language texts. Activities like automatic annotation and content extraction out different images/video/audio could be seen as content extraction.

E. System Calls

In System Calls a programmatic approach takes place in which a computer program requests a service from kernel of the OS it is executed on, which may include functions related to hardware services for example, accessing hard disk drive, creation and execution of new processes and communication with process scheduling. It provides an important interface between the process and the OS.

F. Google-Text-To-Speech

Text-To-Speech is basically used for conversion of Speech from Text provided by the user. In other words, a TTS Engine Converts written form of text into phonemic representation, then converts the phonemic representation to waveforms which results in sound. TTS has developed a lot and comes with different languages provided by the third-party publishers.

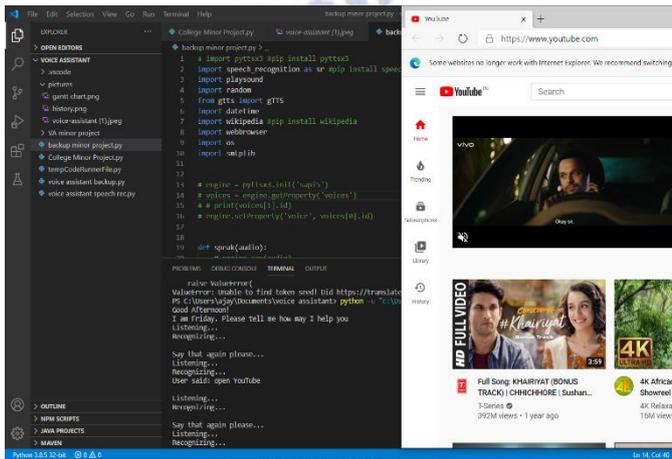


VI. CONCLUSION AND RESULT

In this study, we have developed a voice assistant which can perform any kind of task in exchange of commands given by the users without any error.

We have added more features like it will listen to the users' voice only and will not be activated from environment noise. The modular nature of this project makes it easy to understand and more flexible. We can add more feature without disturbing the whole program and the functionalities. The packages which are required in python programming language has been installed and the code was implemented using VS Code Integrated Development Environment (IDE). The python version used for this project was 3.x and the data of different noises also taken from the environment.

- [8] Tae-Kook Kim “Short Research on Voice Control System Based on Artificial Intelligence Assistant”, Busan, Republic of Korea.



VII. ACKNOWLEDGMENT

This project was done under the guidance of Mr. Bhaskar Kapoor who is also my mentor/guid and supported by department of Information Technology, Maharaja Agrasen Institute of Technology. Thanks for all the support by Bhaskar Kapoor Sir and Maharaja Agrasen Institute of Technology.

REFERENCES

- [1] Subhash S, Prajwal N Srivatsa, Siddesh S, Ullas A, Santhosh B “Artificial Intelligence-Based Voice Assistant” (pp. 593-596). IEEE Conference, (2020).
- [2] Deepak S, Ria U, Monika R, Aishwarya B, Anup B. “AI Based Voice Assistant Using Python” Volume 6 (Issue 2), (pp. 506-509), JETIR, (2019), (ISSN-2349-5162)
- [3] B. Marr, “The Amazing Ways Google Uses Deep Learning AI” (2017).
- [4] Dr. Kulhalli K, Dr. Sirbi K, Mr. Patankar A. (2017). “Personal Assistant with Voice Recognition Intelligence”. Int. Journal of Engineering Research and Technology. Vol. 10 (No. 1) (pp. 416-419), International Research Publication House.
- [5] Kepuska, V., Bohouta, G. “Next generation of Virtual Personal Assistants (Microsoft Cortana, Apple Siri, Amazon Alexa and Google Home)”. IEEE Conference (2018).
- [6] Cortana Intelligence, Google Assistant, Apple Siri.
- [7] G. Bohouta, V. Z. Kēpuska, "Comparing Speech Recognition Systems (Microsoft API Google API)", Int. Journal of Engineering Research and Application (2017).