

# Desktop Voice Assistant

**Jatin Shah and Rini Saxena**

Department of Computer Science and Engineering, Chandigarh Engineering College, Jhanjeri-140307, Punjab, India  
E-mail: rini.j1099@cgce.ac.in

**Abstract:** Artificial intelligence is the growing technology in modern tech science. Artificial Intelligence is a widely-ranging branch of Computer science. The main goal of Artificial intelligence is to enhance the intelligence of machine-like humans so that they can make a decision and Respond to the human's query and do the task that humans can do. The one of prevalent trends in Artificial intelligence is a technology of recognizing the human language. Before this concept of artificial intelligence, we can't automate the machine to do the task as per their own prediction and natural language processing is also not possible. Artificial Intelligence revolutes the industry of computer science. Artificial intelligence makes the machine smart and reduce human work and also helps the other industries to grow. Voice assistant is also one of the Artificial intelligence technology which takes the input as the natural language and responds with a task or query as output. Voice assistants like Google, Siri, Alexa, and Cortana are some voice assistants which make people's life smarter and easier. The natural language processing algorithm is used to process the human being's language to machine understanding language. Now the question is how does it work? How does it takes the natural language as input and recognize it and convert it to machine understanding language and perform the task? Why it's useful?. we will discuss all these in our paper..

## Introduction

A virtual assistant is used to run machines like laptops, PC, and mobile on user command. A virtual assistant application program that takes the natural language as a voice command to perform the task. A virtual assistant can do several tasks like open and close applications or software, play music and video, Set reminders and alarms, control desktop actions, send & read email, read and write documents, search results on the web, send messages on WhatsApp and much more with voice command only. There is some type of voice assistant.

1. Intelligent or Automated personal Assistant
2. Digital Assistant
3. Chatbot

Voice assistant is very useful as it makes human life easy and saves time. An intelligent or Automated voice assistant is a software program that can perform a task on basis of the location awareness or learn the routine habit to perform the daily task automatically or location awareness tasks like weather, traffic, stock price etc. The digital voice assistant is the program software used to do the job of personal assistant or secretary like dictation, reading text, emails, messages, phone calls etc. chatbot is very helpful assistant which is useful for the business application and websites. Before the concept of AI, the business, industries and bank Applications and Websites uses help guides to ease the customer and set up a large call center to help the customer.

The concept of the voice assistant is a revolutionary step in the growth of these industries. Now the chatbot voice assistant is used in these industries' apps and websites to go through the websites and application features and solve the minor issue regarding the products and services to reduce the workload of the

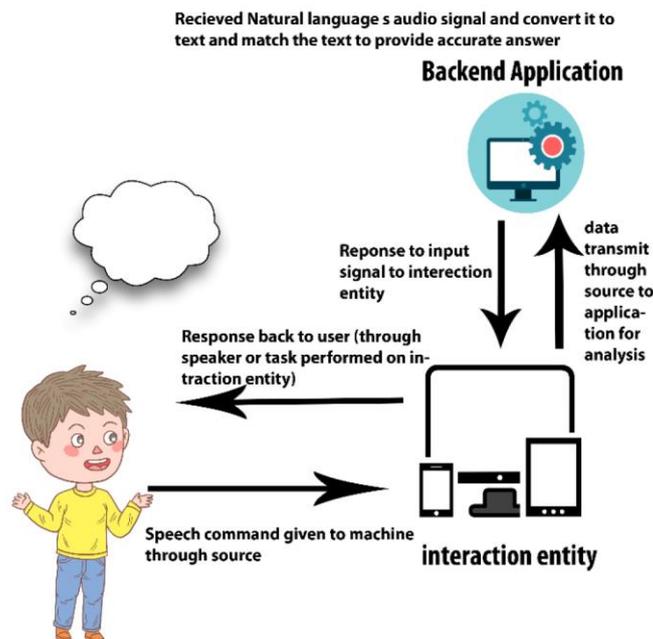
industries. Voice assistants are useful for an individual as well as for industries to grow business. Voice assistant working is the integration of several different technologies like voice recognition, voice analysis and language processing. Voice assistants match the user's natural language command with the function text to perform a specific task or query.

### Problem formulation

Every human has their characteristics and every developer applies his method and approach for developing the product, different developers use different approaches to problem-solving. Some voice assistants can synthesize speech more qualitatively, another can more accurate and another perform a narrow range of tasks. Therefore there is no such assistant who can have all the quality and perform equal action. That depends on the developer and in which area the developer pays more attention. The approach to the problem decides the quality and correctness of the software or product.

### Methodology

The methodology is a process by which we can prepare step to step procedure to develop the best product in the right manner. Virtual assistants use Natural processing language to match a user's text with voice input to execute the tasks. When the user asks the question then the natural language audio signal is converted into an executable command or digital data that can be analyzed by a voice assistant. This Audio is converted into a string and compared with a suitable answer to perform a task. According to this research, the important approach is processing the NLP to text and audio for software to answer the question. Other than this are the additional modules used to add different functionality in our voice assistant. Some of the packages are in Python, which we will discover in this paper. But before this, I want to explore the main packages without which developing a voice assistant would not be possible.



### Speech Recognition as sr

In this python module, it recognizes the natural language audio signal into a string. The speech recognition module uses the Google API for voice recognition as speech.

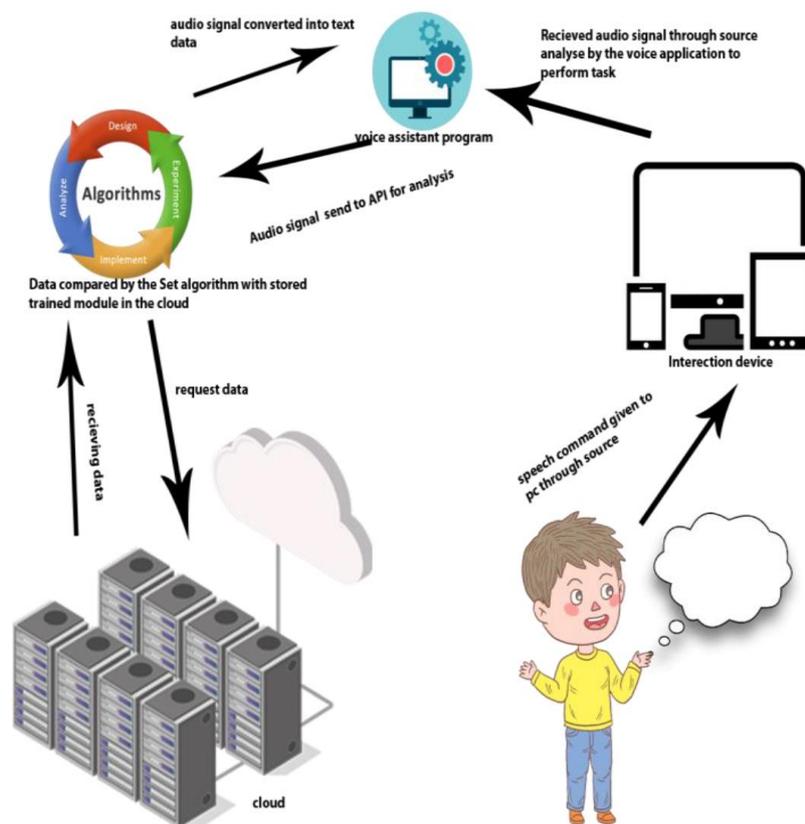
Google API: How does it work?

The concept of the module speech recognition looks very simple from the outside, as we just import the library in Python and use it. But how it gets recognized is still a question for many users. Every step of processing of this process has a different function, algorithm, and trained module which performs this task. Behind this, basically, this is the concept of Deep learning. google used the Neural Network technique to build this API.

Google already has bunches of trained data modules of speech stored in the cloud.

Basically, it using ample cloud processing power, google can analyze a ton of patterns which in voice are spectrograms and use the predict new pattern value like a neuron in the brain would reconnect to complete a task. Google has already stored data of speeches called trained data which use to predict the speech. There is a couple of layer in processing speech. Google firstly tries to understand consonants and vowels. Next, it tries to make an intelligent guess with the help of trained data and go higher.

Eg. It's similar to Lexical analysis in which each syntax or variable, identifier constant, etc is tokenized to compare with the symbol table to get the analyzed machine code. There are special algorithms used in form of technology. It's like a middleware as it is a set of instructions It takes data and matches it with the trained module which is stored in eight houses of the worldwide server of google and combined to predict (like a neuron in the brain) the new value as output. Algorithms are PLP feature, Viterbi search, Deep neural network, discrimination training, WEST framework, etc



Installing speech recognition as sr module, we have to install with this command in terminal  
Pip install speechrecognition as sr

After you also have to install Pyaudio package as it is the package who define the source for taking user voice. If not installed it throw and error microphone not define or undefined.

After installing we just have to import it in python file

Eg.

```
r = sr.recognizer()
with sr.microphone() as source
audio=r.listen(source)
```

### Text-to-speech (pyttsx3 module):

pyttsx3 is text to speech conversion library in python. In this, the text is converted into speech and spoken by the engine already installed in pc. There are 4 voices pre-placed on the PC.

we have to install pyttsx3 from terminal

```
pip install pyttsx3
```

and then import in a python file and define the voice engine :

```
engine= pyttsx.init(here we have to place the TTS engine we want to use like 'sapi5' 'nssss' 'espeak')
```

then we have to get details of current voices by

```
voices=engine.getproperty('voices')
```

if we want to see the voices with voices id we can also print them before we use them

```
print(voices)
```

then we choose one of them by assigning the id of voices in the voices parameter box

```
engine.setProperty('voices',voices[0].id)
```

we have already installed function engine.say(), by which we can give a command to the engine to read the line or whatever we want to read. We have to just put the reading line in the perimeter of engine.say('here').

We can also create the own function to short this method :

```
def speak(audio):
```

```
engine.say(audio)
```

```
engine.runAndWait()
```

speed and volume of voices can also be controlled by

```
engine.setProperty('rate') or speed
```

```
engine.setProperty('volume', 2.0) for volume
```

How Text to speech actually works, the TTS system there are stored phonetic component and sound unit of each word and syntax, by using this inventory the engine extracts them and compare them with the input data and produce the output voices

### Some other Packages which is used to do some special Task

1. Datetime : it is the library in python that allow to work with date and time
2. Wikipedia : this module help in extracting data from the Wikipedia
3. Webbrowser: this module has predefined all browser classes and allow user to do an operation of webbrowser like open search etc. we have to import webbrowser and then we can directly write the command like webbrowser.open(url)
4. OS : This module is preinstalled with python after import this allows to do all operation of the operating system
5. Pywhatkit : this module allow to send and view the message on Web whatsapp
6. Pyjokes: this module is used to extract random jokes in a different language
7. Send Email (smtp) : the SMTP protocol client is used to send email by using Smtplib protocol. We have to create smtp object to use this protocol.  

```
Import smtplib
smtpObj = smtplib.SMTP( [host [, port [, local_hostname]]] )
```

to use this we have to enter the host id and port no usually (587 for Gmail) and then write a function to define the attributes like content, subject, to, etc
8. Pyautogui module: this module allow to perform the operation visible on the screen like drag the mouse and doing mouse operation and keyboard operation and accessing the devices  
we have to import this module and install it with terminal and use there attributes like pyautogui.click(), pyautogui.drag(), pyautogui.press() etc. are some attributes.

**Common Problem:**

Sometimes packages installation throw an error, this is due to some technical issue like python version incompatibility or sometimes unable to fetch data. To overcome this problem try to update the python version or download the .whl file of modules from the python official website and install it directly from the terminal through command `pip install File_path`

2. Problem regarding login in Gmail through smtp module, this is happening because google has stopped the option of the less secure app. To solve this issue we have to generate a hash code from 2 step verification factor from Gmail and enter this in place of the password  
`server.login('@gmail.com', 'hashcode')`

3. Recognition problems like listening when not speaking, we can solve this by adding the property of `“recognizer_instance.energy_threshold property”`

**Result**

Voice assistant is less time-consuming and does the task quickly on voice command by the user. It uses NLP to match the voices of the user or text input to give the desired output. With the help of the voice assistant, we can control pc with our voice and also send messages and emails and do searches. Voice assistant makes human life easier. Voice assistant is the need of the future. This is not only helpful for individuals but also for the industries. It helps them to grow and save resources and time like chatbot voice assistants assist the customer. Thus its saves time for industries and resources and manpower. Even a blind person can also operate a PC, enjoy music send emails read documents and emails, and so on.

**Conclusion**

In this paper, we have discussed the working of Voice assistants and their benefits and uses. How its works and gives desired output. We can also make the voice assistant in python for all window versions. It is basically the concept of artificial intelligence. We can also add an Interface by using the Tkinter module or pyqt5. Voice assistant is the need of the future and everyone is adopting this as a feature of his app or website. I am attaching the link to the code of the voice assistant made by me for others references.

**References**

- [1] DOUGLAS O'SHAUGHNESSY, SENIOR MEMBER, IEEE, “Interacting With Computers by Voice: Automatic Speech Recognition and Synthesis” proceedings of THE IEEE, VOL. 91, NO. 9, SEPTEMBER 2003
- [2] Kei Hashimoto<sup>1</sup>, Junichi Yamagishi<sup>2</sup>, William Byrne<sup>3</sup>, Simon King<sup>2</sup>, Keiichi Tokuda, “An analysis of machine translation and speech synthesis in speech-to-speech translation system” proceedings of 5108978-1-4577-0539-7/11/\$26.00 ©2011 IEEE.
- [3] Nil Goksel-Canbek<sup>2</sup> Mehmet Emin Mutlu, “On the track of Artificial Intelligence: Learning with Intelligent Personal Assistant” International Journal of Human Sciences, 13(1), 592-601. doi:10.14687/ijhs.v13i1.3549.
- [4] Harshita Phatnani, Mr. Jyotiprakash Patra and Ankit Sharma’ “CHATBOT ASSISTING: SIRI” Proceedings of BITCON-2015 Innovations For National Development National Conference on : Research and Development in Computer Science and Applications, E-ISSN2249–8974.
- [5] Sutar Shekhar, Pophali Sameer, Kamad Neha, Prof. Devkate Laxman, " An Intelligent Voice Assistant Using Android Platform" , March 2015, IJARCSMS, ISSN: 232 7782
- [6] VINAY SAGAR, KUSUMA SM, "Home Automation Using Internet of Things", June-2015, IRJET, e-ISSN: 2395 -0056
- [7] <https://www.cnet.com/products/ivee-sleek/reviews/>
- [8] <https://www.cnet.com/products/athom-homey/ ity>