

# Artificial Intelligence Based A Communicative Virtual Voice Assistant Using Python & Visual Code Technology

Raj Kumar Jain, Vikas Sharma, Mangilal, Rakesh Kardam, Mamta Rani

**Abstract**— The advancements in technologies over the time have been un-measurable. From the first digital personal computer built by Mr. Eniac having a clock speed of 0.1 MHz to summit developed by the United States Department of Energy has a performance of 148.6 peta flops, and we have come a very long way in technological advancement. In this era of advancement if people are still struggling to interact with their machine using various input devices then it's not worth it. Because of this reason, many voice assistants have developed and still are being improved for better performance and efficiency.

The main work of a voice assistant prototype is to minimize the usage of input devices like touch pens, keyboard, mouse, etc. This will reduce the hardware cost and space taken by it both.

**Index Terms**— Desktop Assistant, Python, Machine Learning, Text to Speech, Speech to Text, Language Processing, Voice Recognition, Artificial Intelligence, Internet Of Things (IOT), Virtual Assistant.

## I. INTRODUCTION

In the 21st century, human interaction is being replaced by automation very quickly. One of the main reasons for this change is performance. There's a drastic change in technology rather than advancement. In today's time, we train our machines to do the tasks by themselves or to think like human being using technologies like Machine Learning, Neural Networks, etc. Now in the current era, we can talk to our machines with the help of virtual assistants. There are companies like Google, Apple, Microsoft, etc with virtual assistants like Google Now, Siri, Cortana, etc. which helps the users to control their machines by just giving input in the form of their voice.

These types of virtual assistants are very useful for old age, blind & physically challenged people, children, etc. by making sure that the interaction with the machine is not a challenge anymore for people. Even blind people who couldn't see the machine can interact with it using their voice only.

Some of the tasks that are supported by most of the virtual assistants (VA) are:

- Checking weather updates
- Sending and checking mails
- Search on Wikipedia
- Make and receive calls

Raj Kumar Jain, Vikas Sharma, Mangilal, Rakesh Kardam, Mamta Rani., Department of Electronics & Communication, Jaipur Engineering College and Research Centre, Jaipur (Raj)

- Stream music
- Open applications
- Text messages etc.

The voice assistant we have proposed and developed is a desktop-based built using python language modules and libraries. This voice assistant is just a basic version that could do all the basic tasks which have been mentioned as above but current technology is although good in it is still to be merged with Machine Learning language (ML) and Internet of Things (IOT) for better enhancements

The understanding and executing commands are still to reach a new level like the virtual assistant of the iron man named Jarvis. This is although fictional yet this is what that can be achieved using virtual assistants. All you need to do is give a command to the assistant and the rest will be performed by the assistant.

With the help of voice-activated virtual assistants, there will be no need to write long codes to perform a task, the system will do so for us. The machine will work in three modes- supervised, unsupervised or reinforcement learning depending upon the usage for which the assistant is developed. This is all possible with the help of machine learning.

Now what the IOT does is it'll help the assistant to interact with the neighboring smart devices and can act as one interface which will control everything within the surrounding. With the involvement of IOT, it'll be possible to regulate other smart devices which will in-turn interact among themselves over the web.

So with a capable virtual assistant, we will be able to control many things around us single-handedly with only one platform



Figure 1: Timeline Of Main Voice Assistants

## II. LITERATURE SURVEY

This field of virtual assistants (VA) having speech recognition have seen some major advancements and innovations. This is because of its big demand in devices viz. smartwatches or fitness bands, smart speakers, Bluetooth earphones, smart mobile phones, AI based laptop or desktop, television, etc.

Almost all the digital devices which are coming nowadays are coming with voice assistants which help to regulate the device with speech recognition only. A new set of AI techniques is being developed continuously to improve the performance of voice automated search. As the amount of knowledge is increasing exponentially now referred to as Big Data the simplest thanks to improve the results of virtual assistants is to include our assistants with machine learning and train our devices consistent with their uses. Other new techniques that are most and equally important are Artificial Intelligence, Internet Of Things, Big Data access etc.

With the utilization of voice assistants, we will automate the task easily, just give the input to the machine within the speech form and every one the tasks are going to be done by this, from converting voice speech into text form to taking out keywords from that text and execute the query to give results to the user.

Machine Learning programming is just a subset of Artificial Intelligence technique. This has been one of the most important and helpful advancements in this technology. Before AI we were the ones, who upgrading constantly technology to compete a task but nowadays the machine is itself able to counter defend new tasks and can solve it without need of involving the human to evolve it.

This has been helpful in day-to-day lifestyle nowadays. From smart mobile phones to personal desktops to automation and mechanical industries these kind of assistants are in very much demand to do automating tasks and increasing efficiency.



Figure 2: Voice Controlled Appliances Affecting Our Daily Life

## III. SYSTEM STRUCTURE

```
import speech_recognition as sr
import os
import sys
import re
import webbrowser
import smtplib
import requests
```

Fig.3: Modules Imported

### • SPEECH RECOGNITION

The speech recognition module used the program is Google's Speech Recognition API which is imported in python using the command "import speech\_recognition as sr". This module is employed to acknowledge the voice which is given as input by the user.

This is a free API that's provided and supported by Google. This is a really light API that helps in reducing the dimensions of our application.

### • TTS & STT

The voice which is given as input is first converted to text using the speech recognition module. The text is then processed to give the result of the query given by the user. The final step is the conversion of the result of the processed query to speech which is the final output. The most time consuming among the two is STT because the system first has to listen to the user and different users have different, some are easy to understand while some are not easily audible. This is the step upon which our total execution time depends. Once the speech is converted in text, executing commands giving the result back to user is not a time consuming step.

### • IMPORTED MODULES

#### i. PYTTX3

The pyttx3 is an offline library module that is used for text to speech conversion in Python language and it is usually supported by both Python 2 & 3. The run and wait functionality is also in this module only. It determines for how much time the system would wait for the another input or in other words the time interval between applied inputs. This is a free module available in the python community which can be installed using the pip command just like other modules.

#### ii. DATETIME

The Date Time module is usually imported to support the functionality of the time and date. Like, if user wants to know the current latest date and time or the user would like to schedule a task at a certain time. In short this is a module which supports classes to manipulate or alter date and time and perform the operations according to it only.

#### iii. WEB BROWSER

This module allows the system to show and display web based information to the users. For example, the user wishes to open a particular website and he gives the input as "Open Google". The given voice input is pre processed using the web browser module and the user will get a browser with google opened in it. The web browser which will be used is always the default set web browser.

#### iv. WIKIPEDIA

Wikipedia, it is a library in python language which it possible for the virtual assistant (VA) to process the queries in regard of Wikipedia and displays the results to the users. This is an online available library and it needs an internet connection to fetch the details and results.

The number of lines that user wants to get as result can be set manually.

#### v. OS MODULE

The Operating System Module provides an operating system dependent functionalities in the python. If we want to execute operations on files like writing, reading or manipulating paths, all these kind of functionalities are easily available in an OS module. All the operations available generally raise an error "OSError" in the case of any error viz. invalid names, paths, or any other arguments which may be incorrect or may be correct but just no accepted by the operating system.

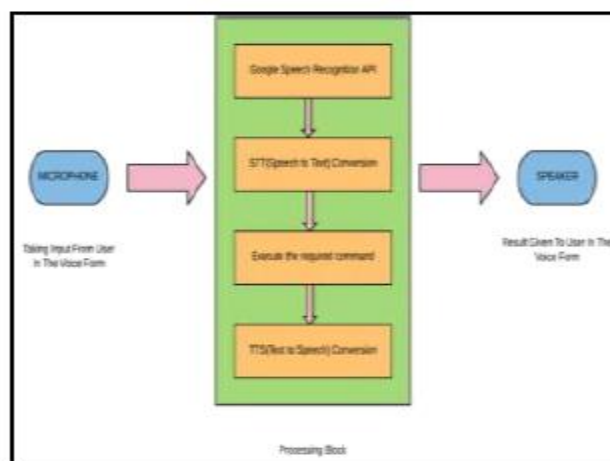
#### vi. SMTPLIB

The Python programming has this module in the standard library for dealing with emails & email servers. The SMTPLIB module defines an object which known as "SMTP client session object", it is used to send mails by the user. There are basically 3 steps involved - initialize, sendmail() and quit. When all the optional parameters which are host and port, are provided connect method is called with these arguments during the first step which named is initialization.

### IV. DESIGN

The overall design of the system consists of the subsequent phases:

- a) Receiving input from the user in the form of voice input
- (b) Conversion of the the speech into text and to be processed by the assistant.
- (c) The converted text signal is processed to get the desired results.
- (d) The text contains one or more keywords that determine the query is to be executed. If the received keyword doesnot match any of the queries in the designed code then the voice assistant asks the user to speak again.
- (e) The results which are in the form of text are converted to speech signals again to give results to the user.



A. Figure 3: Processing Block of STT To TTS

### V. PROPOSED SYSTEM

The proposed system would have the following functionality:

- (a)The system will be keep listening for the commands and time for listening is variable which can be changed according to the user requirements.
- (b)In any case, If the system is not able to gather information from the user input it will keep asking again to repeat till the desired no. of times.
- (c)The system could have male and female both voices according to the users requirements.
- (d)Many Features supported in the proposed current version including playing music songs, accessing emails, texts messages, searching on Wikipedia, or opening and access of system installed applications, opening anything on the web browser, etc.
- (e)The system will keep listening for the commands and time for the listening is can be variable which changed according to user requirements.
- (f)If the system is not able to gather the information from user input, it will keep asking again and again to repeat till the desired no. of times.

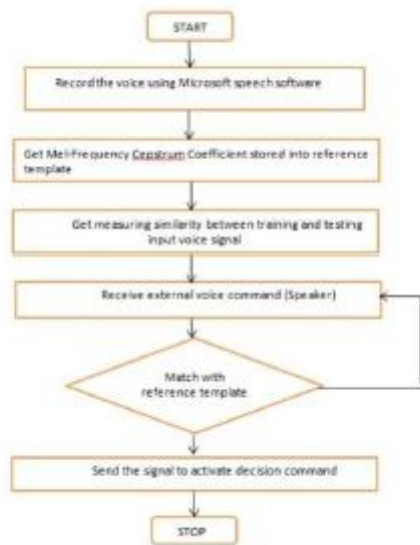


Fig7: Flowchart for Voice Flow Algorithm

### VI. FUTURE SCOPE

The virtual assistants which are currently available are fast and responsive but we still need to go an extended way. The understanding and reliability of the present systems got to be improved tons .

The assistants available nowadays are still not reliable in critical scenarios. The future of those assistants will have the virtual assistants incorporated with AI which incorporates Machine Learning, Neural Networks, etc. and IoT.

With the incorporation of those technologies, we'll be ready to achieve new heights. What the virtual assistants can achieve is much beyond what we have achieved till now. Most folks have seen Jarvis, that's a virtual assistant developed by iron man which is although fictional but this has set new standards of what we will achieve using voice-activated virtual assistants.

### VII. CONCLUSION

In this paper we have discussed Jarvis Based AI Desktop Voice Assistant using python. This assistant currently works online and performs basic tasks like weather updates, stream music, search Wikipedia, open desktop applications, etc.

The functionality of the current system is limited to working online only. The upcoming updates of this assistant will have machine learning incorporated in the system which will result in better suggestions with IoT to control the nearby devices similar to what Amazon’s Alexa does.

The usage of the assistant will get offline also for features that don’t require an internet connection.

### REFERENCES

[1] D 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, Junichi Yamagishi, William Byrne, Simon King, 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 Mehmet Emin Mutlu, “On the track of Artificial Intelligence: Learning with Intelligent Personal Assistant” International Journal of Human Sciences.

[4] H. Phatnani, Mr. J. 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, P. 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] “Speech recognition with flat direct models,” IEEE Journal of Selected Topics in Signal Processing, 2010.

[8] Rishabh Shah, Siddhant Lahoti, Prof. Lavanya. K, “An Intelligent Chatbot using Natural Language Processing”. International Journal of Engineering Research, Vol. 6, pp. 281- 286, 1 May 2017.