iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 10 **Issue:** XII **Month of publication:** December 2022

DOI: https://doi.org/10.22214/ijraset.2022.48067

www.ijraset.com

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

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 10 Issue XII Dec 2022- Available at www.ijraset.com

Voice Activated Desktop Assistant Using Python

Indranil Basu[1], Saumyadeep Bhattacharyya[2], Arpan Mondal[3], Babin Maitra[4], Roshni Joardar[5], Toyesh Dey[6], Koushik Pal[7]

[1, 2, 3 ,4, 5, 6, 7]*Electronics and Communication Engineering Department, Guru Nanak Institute of Technology, Kolkata*

Abstract: ***The development in science over time has been unmeasurable. From the first digital laptop constructed by using Eniac having a clock pace of 100KHz to Summit developed by the US Department of Energy has an overall performance of 148.6 petaflops, we have come a lengthy way in technological advancement. In such a generation of development if human beings are still struggling to interact with their laptop the usage of a variety of input units, then it's no longer worth it. For this reason, many voice assistants had been developed and are nonetheless being accelerated for higher overall performance and efficiency. The fundamental mission of a voice assistant is to limit the use of input units like keyboard, mouse, contact pens, etc. This will limit each the hardware price and house taken with the aid of it.***

Keywords: 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 twenty first century, human interplay is being changed by way of automation very quickly. One of the predominant motives for this change is performance. There's a drastic trade in technological know-how rather than advancement. In today's world, we instruct our machines to do their duties via themselves or to assume like humans the use of applied sciences like Machine Learning, Neural Networks, etc. Now in the cutting-edge era, we can speak to our machines with the assist of digital assistants. There are companies like Google, Apple, Microsoft, and so on with digital assistants like Google Now, Siri, Cortana, etc. which helps their users to manipulate their laptop by means of simply giving enter in the shape of voice.

These sorts of digital assistants are very beneficial for ancient age, blind & bodily challenged people, children, etc. via making sure that the interplay with the computer is no longer an undertaking anymore for people. Even blind human beings who couldn't see the machine can interact with it the usage of their voice only.

Some of the primary duties that are supported via most of the digital assistants are:

1) Checking climate updates
2) Sending and checking mails
3) Search on Wikipedia
4) Make and get hold of calls
5) Stream music
6) Open applications
7) Text messages etc.

The voice assistant we have developed is a desktop-based constructed using python modules and libraries. This assistant is simply a fundamental version that may want to operate all the simple duties which have been mentioned above but modern science is even though excellent in it is nevertheless to be merged with Machine Learning and Internet Of Things (IoT) for higher enhancements.

The appreciation and executing instructions are nevertheless to attain a new degree like the digital assistant of the iron man named Jarvis. This is even though fictional, but this is what that can be accomplished using digital assistants. All you want to do is provide a command to the assistant and the relaxation will be carried out by means of the assistant. With the assist of voice-activated digital assistants, there will be no want to write lengthy codes to operate a task, the device will do so for us. The desktop will work in three modes- supervised, unsupervised or reinforcement mastering relying upon the usage for which the assistant is developed. This is all viable with the assist of laptop learning.

Now what the IoT does is it will assist the assistant to interact with the neighbouring clever gadgets and will act as a single interface that will manipulate the whole thing in the surrounding. With the involvement of IoT, it will be viable to manipulate different clever devices that will in-turn engage amongst themselves over the internet.

So, with a successful digital assistant, we will be capable to manage many matters round us single-handedly with solely one platform

©IJRASET: All Rights are Reserved

Fig. 1 Timeline of Main Voice Assistants

II. LITERATURE SURVEY

This subject of digital assistants having speech attention has seen some foremost developments or innovations. This is usually because of its demand in gadgets like smartwatches or health bands, speakers, Bluetooth earphones, cell phones, laptop computer or desktop, television, etc. Almost all the digital units which are coming presently are coming with voice assistants which assist to manipulate the machine with speech focus only. A new set of techniques is being developed continuously to enhance the performance of voice automatic search. As the quantity of information is growing exponentially now recognized as Big Data the exceptional way to enhance the consequences of digital assistants is to include our assistants with computer studying and instruct our gadgets in accordance with their uses. Other main techniques that are equally necessary are Artificial Intelligence, Internet of Things, Big Data get admission to and management, etc. With the use of voice assistants, we can automate the assignment easily, just provide the enter to the computing device in the speech structure and all the tasks will be completed with the aid of it from changing your speech into textual content form to taking out key phrases from that textual content and execute the query to provide consequences to the user. Machine Learning is simply a subset of Artificial Intelligence. This has been one of the most beneficial developments in technology. Before AI we have been the ones who have been upgrading technological know-how to do an assignment however now the computer is itself capable to counter new duties and clear up it besides want to contain the human beings to evolve it.

This has been beneficial in daily lifestyle. From cell phones to private pcs to mechanical industries these assistants are in very plenty demand for automating duties and increasing efficiency.

Fig. 2 Voice Controlled Appliances Affecting Our DailyLife

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 10 Issue XII Dec 2022- Available at www.ijraset.com

III. SYSTEM ARCHITECTURE

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

Fig. 3 Modules Imported

A. *Speech Recognition*

The speech attention module used the software is Google's Speech Recognition API which is imported in python the use of the command "import speech recognition as sr". This module is used to understand the voice which is given as enter via the user. This is a free API that is supplied and supported through Google. This is a very mild API that helps in lowering the dimension of our application.

B. *TTS & STT*

The voice which is given as enter is first transformed to textual content the usage of the speech awareness module. The textual content is then processed to supply the result of the question given through the user. The ultimate step is the conversion of the result of the processed question to speech which is the ultimate output. The most time eating amongst the two is STT due to the fact the device first must pay attention to the person and distinct customers have different, some are handy to apprehend whilst some are now not without problems audible. This is the step upon which our whole execution time depends. Once the speech is transformed to textual content executing instructions and giving the consequences returned to the person is no longer a time-consuming step.

C. *Imported Modules*

1) *PYTTSX3:* The pyttsx3 is an offline module that is used for textual content to speech conversion in Python and it is supported via both Python two & three. The run and wait performance is additionally in this module solely. It determines how a lot of time the device will wait for any other enter or in different phrases the time interval between inputs. This is a free module accessible in the python neighbourhood which can be hooked up the usage of the pip command simply like different modules.
2) *Date Time:* The Date Time module is imported to aid the performance of the date and time. For example, the person desires to comprehend the modern date and time or the consumer desires to agenda a mission at a positive time. In brief this module helps lessons to manipulate date and time and function operations in accordance with it only. This is a quintessential module, in duties the place we choose to hold a song of time. This module is very small in measurement and helps to manage the dimension of our program. If the modules are too massive or heavy, then the device will lag and provide gradual responses.
3) *Web Browser:* This module lets in the gadget to show web-based facts to users. For example, the consumer desires to open any internet site and he offers enter as "Open Google". The enter is processed the use of the net browser module and the consumer receives a browser with google opened in it. The browser which will be used is the default set net browser.
4) *WikiPedia:* Wikipedia is a library in python which it viable for the virtual assistant to system the queries concerning Wikipedia and show the outcomes to users. This is an online library and desires a web connection to fetch the results. The no. of traces that the consumer needs to get as a result can be set manually.
5) *OS Module:* OS Module presents a working device established functionalities. If we favour to operate operations on archives like reading, writing, or manipulate paths, all these kinds of functionalities are accessible in an OS module. All the operations handy increase an error "OS Error" in case of any error like invalid names, paths, or arguments which might also be mistaken or right however simply no ordinary with the aid of the running system.
6) *SMTPLIB:* Python has this module for in the general library for working with emails & e mail servers. The SMTPLIB defines an object acknowledged as "SMTP consumer session object" which is used to send mails by using the user. There are three steps worried - initialize, sendmail(), quit. When the non-obligatory parameters which are host and port, are supplied join technique is known as with these arguments throughout the first step which is initialization.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

D. Design

The typical diagram of our gadget consists of the following phases:

1) Taking enter from the consumer in the structure of voice.
2) Converting the speech into textual content to be processed by using the assistant.
3) The transformed textual content is now processed to get the required results.
4) The textual content carries one or two key phrases that decide what question is to be executed. If the keyword doesn't suit any of the queries in the code, then the assistant asks the consumer to talk again.
5) The result which is in the structure of textual content is transformed to speech once more to supply effects to the user.

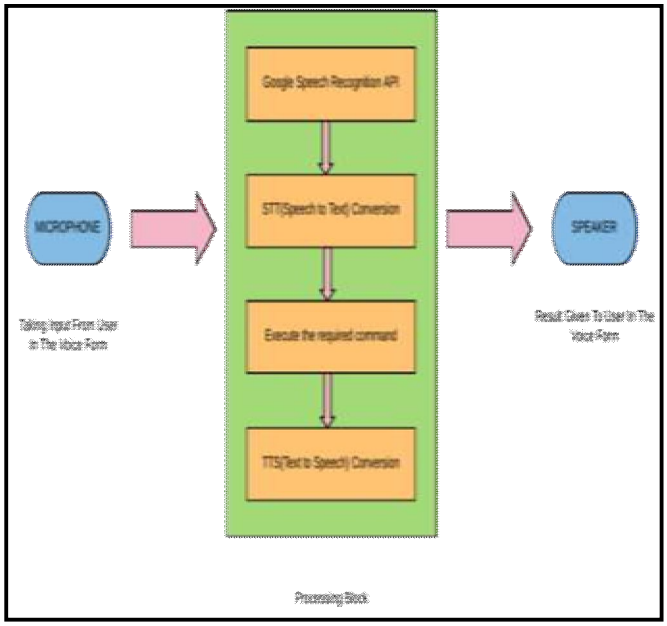


FIG. 4 PROCESSING BLOCK OF STT TO TTS

E. Proposed System

The proposed device will have the following functionality:

1) The gadget will hold listening for instructions and the time for listening is variable which can be modified in accordance with consumer requirements.
2) If the device is no longer capable to acquire statistics from the consumer enter it will preserve asking once more to repeat until the favoured no. of times.
3) The gadget can have each male and girl voices in accordance with consumer requirements.
4) Features supported in the present-day model encompass taking part in music, emails, texts, search on Wikipedia, or opening device established applications, opening something on the net browser, etc.
5) The machine will hold listening for instructions and the time for listening is variable which can be changed in accordance with person requirements.
6) If the gadget is now not capable to accumulate data from the consumer enter it will hold asking once more to repeat until the favoured no. of times.
7) The gadget can have each male and woman voices in accordance with consumer requirements.

F. Future Scope

The digital assistants which are presently accessible are speedy and responsive however we nevertheless must go a lengthy way. The appreciation and reliability of the modern structures want to be increased a lot. The assistants on hand currently are nevertheless no longer dependable in imperative scenarios. The future of these assistants will have the digital assistants integrated with Artificial Intelligence which consists of Machine Learning, Neural Networks, etc. and IoT. With the incorporation of these technologies, we will be capable to gain new heights. What the digital assistants can gain is plenty past what we have executed until now. Most of us have viewed Jarvis, that is a digital assistant developed by means of iron man which is even though fictional however this has set new requirements of what we can obtain the usage of voice-activated digital assistants.

©IJRASET: All Rights are Reserved

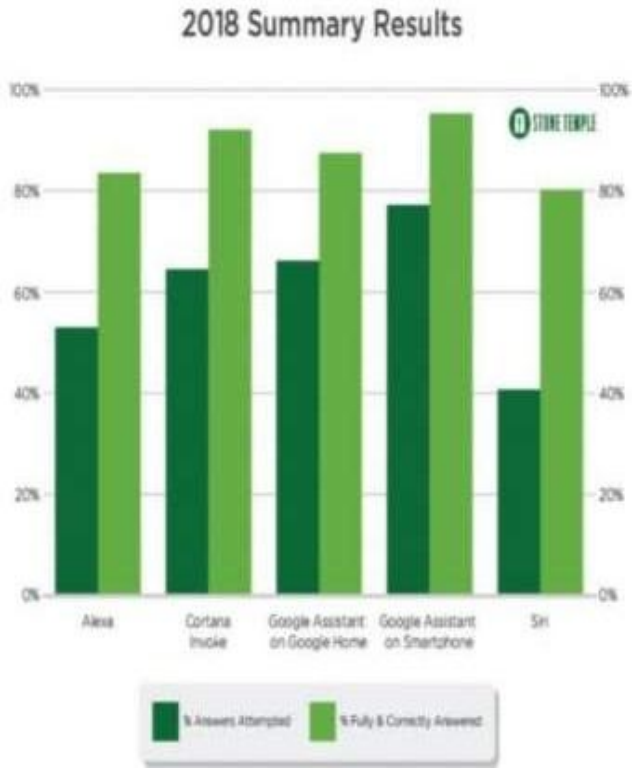


Fig. 5 Accuracy of Results Over Time

IV. CONCLUSIONS

In this paper we have mentioned a Voice Activated Personal Assistant developed the usage of python. This assistant presently works on-line and performs fundamental duties like climate updates, circulate music, search Wikipedia, open computer applications, etc. The performance of the present-day machine is restrained to working on-line only. The upcoming updates of this assistant will have computer mastering included in the machine which will result in higher recommendations with IoT to manipulate the close by units comparable to what Amazon's Alexa does

The utilization of the assistant will get offline additionally for elements that don't require a net 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.

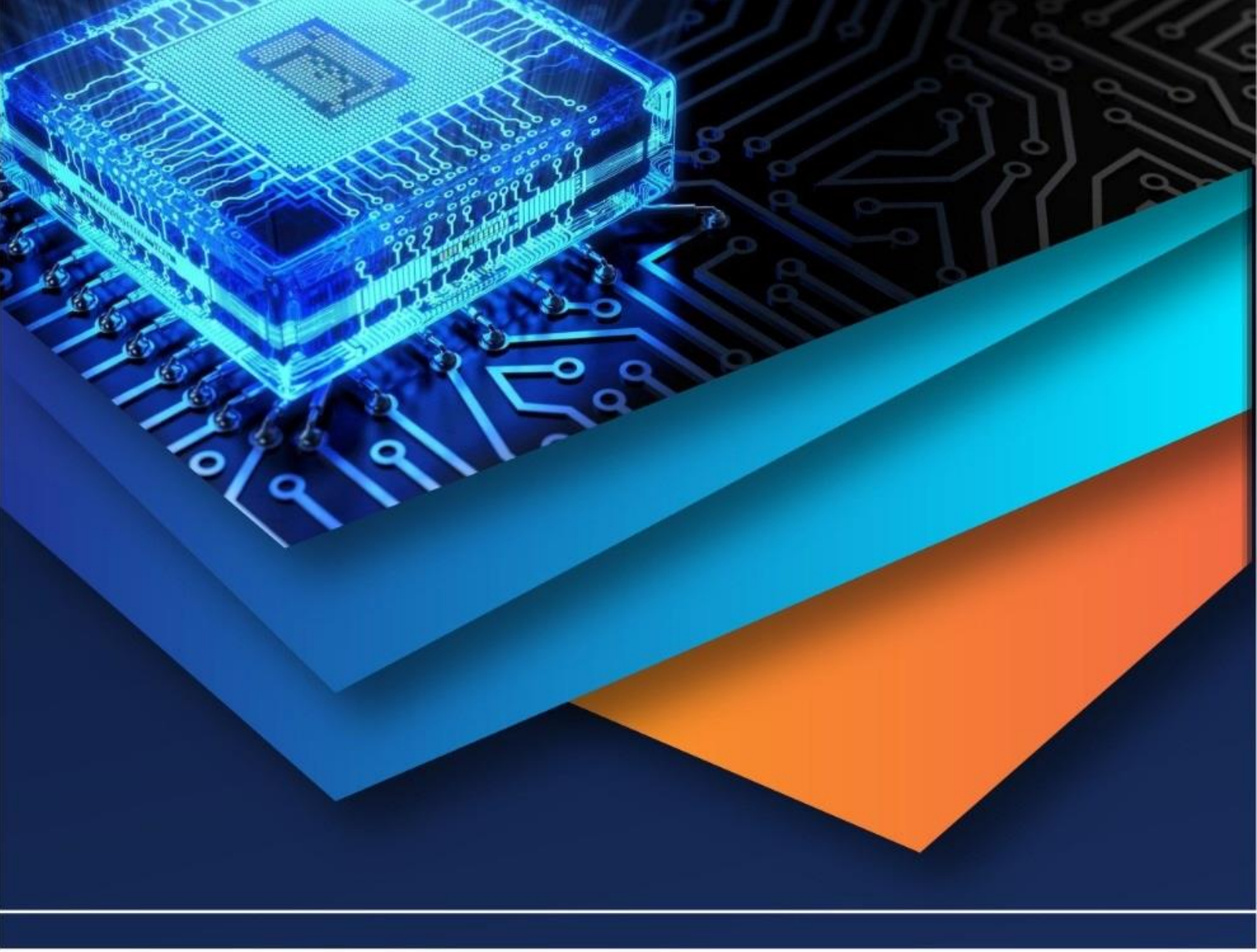
[4] H. Phatnani, Mr. J. Patra and Ankit Sharma' "CHATBOT ASSISTING: SIRI" Proceedings of BITCON-2015 Innovations For National Development National Conference on Research and Development in Computer Science and Applications, E-ISSN2249–8974.

[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.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)