IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Vol. 10, Issue 10, October 2021

DOI: 10.17148/IJARCCE.2021.101015

DESKTOP SPEECH RECOGNITION

Pavithra K[1], Pooja Sree V[2], Priyadarshini S[3]

[1,2,3]UG Scholar Department of Computer Science , Paavai Engineering College Autonomous , Namakkal.

Abstract: Speech recognition technology is one from the fast growing engineering technologies. It has a number of applications in different areas and provides potential benefits. Nearly 20% people of the world are suffering from various disabilities; many of them are blind or unable to use their hands effectively. The speech recognition systems in those particular cases provide a significant help to them, so that they can share information with people by operating computer through voice input.

This project is designed and developed keeping thatfactorinto mind, and a little effort is made to achieve this aim. Our project is capable to recognize the speech and convert the input audio into text; it also enables a user to perform operations such as "save, open ,exit" a file by providing voice input. It also helps the user to open different system software such as opening notepad, music player, you tube, etc.

At the initial level effort is made to provide help for basic operations discussed above, but the software can further be updated and enhanced in order to cover more operations.

INTRODUCTION:

A personal voice assistant is the software that can perform task and provide different services to the individual as per the individual's dictated commands. This is done through a synchronous process involving recognition of speech patterns and then, responding via synthetic speech. Through these assistants a user can automate tasks ranging from but not limited to mailing, tasks management and media playback. As the technology is developing day by day people are becoming more dependent on it, one of the mostly used platform is computer. We all want to make the use of these computers more comfortable, traditional way to give a command to the computer is through keyboard but a more convenient way is to input the command through voice. Giving input through voice is not only beneficial for the normal people but also for those who are visually impaired who are not able to give the input by using a keyboard. For this purpose, there is a need of a voice assistant which can not only take command through voice but also execute the desired instructions and give output either in the form of voice or any other means.

Keywords: Python script, speech recognition, voice assistant Abbreviation: API (Application program interface), NLP (natural language processing), TTS (Text-To-Speech).

LITERATION SURVEY:

This field of virtual assistants having speech recognition has seen some major advancements or innovations. This is mainly because of its demand in devices like smartwatches or fitness bands, speakers, blue tooth earphones, mobile phones, laptop or desktop, television, etc.

Almost all the digital devices which are coming nowadays are coming with voice assistants which help to control the device with speech recognition only. A new set of techniques is being developed constantly to improve the performance of voice automated search

As the amount of data is increasing exponentially now known as Big Data the best way to improve the results of virtual assistants is to incorporate our assistants with machine learning and train our devices according to their uses. Other major techniques that are equally important are Artificial Intelligence, Internet of Things, Big Data access and management, etc.

With the use of voice assistants, we can automate the task easily, just give the input to the machine in the speech form and all the tasks will be done by it from converting your speech into text form to taking out keywords from that text and execute the query to give results to the user. Machine Learning is just a subset of Artificial Intelligence.

This has been one of the most helpful advancements in technology. Before AI we were the ones who were upgrading technology to do a task but now the machine is itself able to counter new tasks and solve it without need to involve the humans to evolve it. This has been helpful in day-to-day lifestyle. From mobile phones to personal desktops to mechanical industries these assistants are in very much demand for automating tasks and increasing efficiency.

RELATED WORK:

Each company developer of the intelligent assistant applies his own specific methods and approaches for development, which in turn affects the final product. One assistant can synthesize speech more qualitatively, another can more accurately and without additional explanations and corrections perform tasks, others can perform a narrower range of

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Vol. 10, Issue 10, October 2021

DOI: 10.17148/IJARCCE.2021.101015

tasks, but most accurately and as the user wants. Obviously, there is no universal assistant who would perform all tasks equally well. The set of characteristics that an assistant has depends entirely on which area the developer has paid more attention to. Since all systems are based on machine learning methods and use for their creation huge amounts of data collected from various sources and then trained on them, an important role is played by the source of this data, be it search systems, various information sources or social networks. The amount of information from different sources determines the nature of the assistant, which can result as a result. Despite the different approaches to learning, different algorithms and techniques, the principle of building such systems remain approximately the same. Figure 1 shows the technologies that are used to create intelligent systems of interaction with a human by his natural language. The main technologies are voice activation, automatic speech recognition, Teach-To-Speech, voice biometrics, dialogue manager, natural language understanding and named entity recognition.

PROPOSED WORK:

1. Methodology:

As an emerging technology, not all developers are familiar with speech recognition technology. While the basic functions of both speech synthesis and speech recognition takes only few minutes to understand, there are subtle and powerful capabilities provided by computerized speech that developers will want to understand and utilize.

Despite very substantial investment in speech technology research over the last 40years, speech synthesis and speech recognition technologies still have significant limitations. Most importantly, speech technology does not always meet the high expectations of users familiar with natural human-to-human speech communication.

An understanding of the capabilities and limitation of speech technology is also important for developers in making decisions about whether a particular application will benefit from the use of speech input and output.

2. Proposed Architecture:

The system design consists of

1. Taking the input as speech patterns through microphone.
2. Audio data recognition and conversion into text.
3. Comparing the input with predefined commands
4. Giving the desired output

The initial phase includes the data being taken in as speech patterns from the microphone.in the second phase the collected data is worked over and transformed into textual data using NLP. In the next step, this resulting signified data is manipulated through Python Script to finalize the required output process. In the last phase, the produced output is presented either in the form of text or converted from text to speech using TTS.

3. Features:

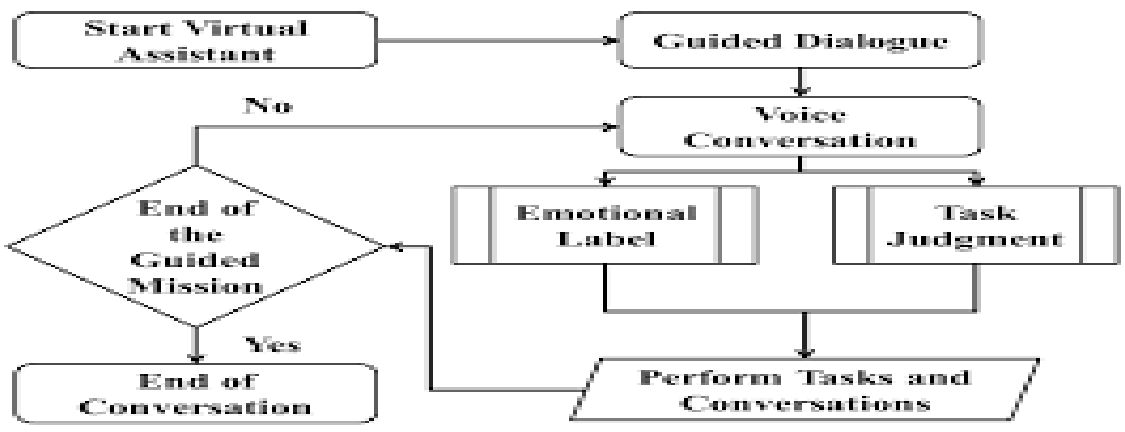
The System shall be developed to offer the following features:

1) It keeps listening continuously in inaction and wakes up into action when called with a particular predetermined functionality.

2) Browsing through the web based on the individual's spoken parameters and then issuing a desired output through audio and at the same time it will print the output on the screen.

Other useful services such as playing any kind of media, browsing weather forecasts, setting, reminders, shut down computer, sending an Email etc. Are provided as a result of spoken commands.

Flow diagram:



IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Vol. 10, Issue 10, October 2021

DOI: 10.17148/IJARCCE.2021.101015

STEP1:

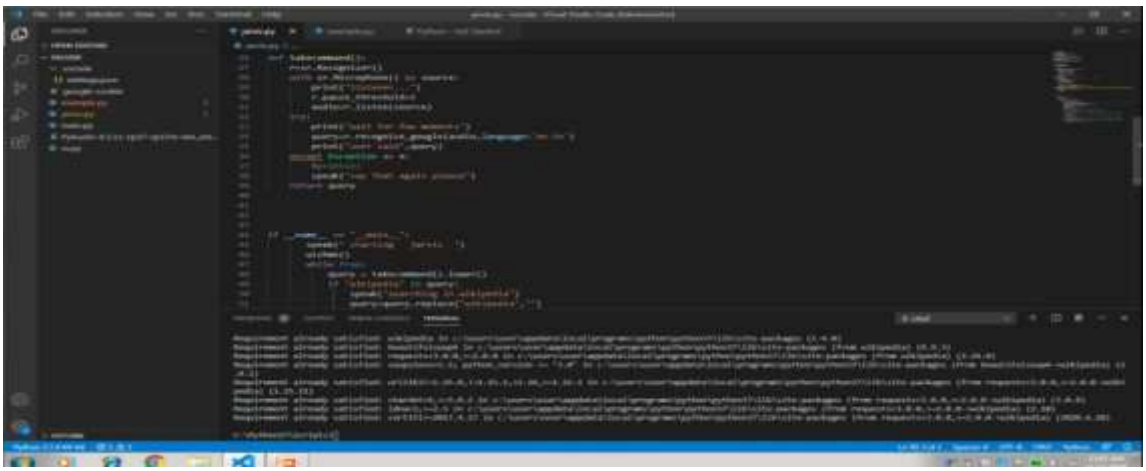
(import packages for this program):

- Import pyttxs3
- Import datetime
- Import speech_recognition as sr
- Import Wikipedia
- From googlesearch import search
- From youtube_search import youtubesearch
- Import urlib.request
- Import os
- Import webbrowser

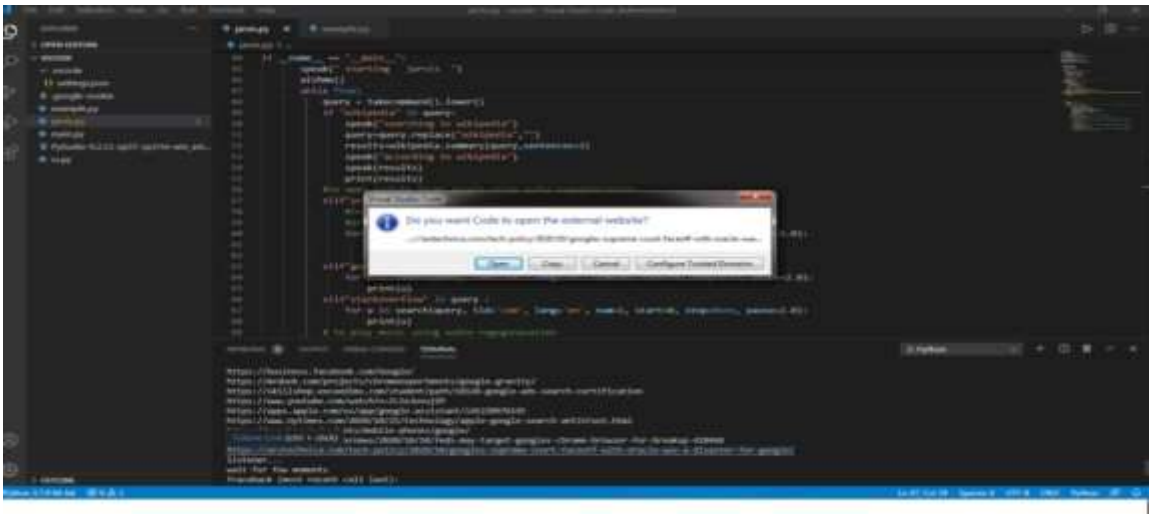
PIP INSTALLATION IN THIS PROGRAM WHICH IS ESSENTIAL:

- PIP INSTALL PYLINT
- PIP INSTALL PYTTSX3
- PIP INSTALL SPEECH RECOGNITION
- PIP INSTALL WIKIPEDIA
- PIP INSTALL GOOGLE
- PIP INSTALL YOUTUBE-SEARCH
- PIP INSTALL WEBBROWSER

Step2: **Initial definition of functions to call**



Step3: **Run the program**



IJARCCE

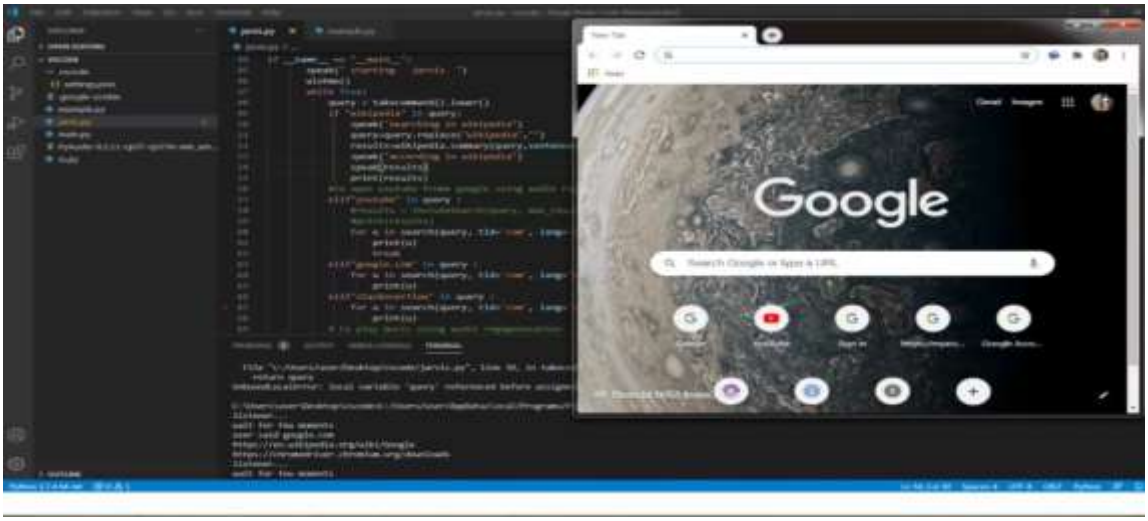
ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Vol. 10, Issue 10, October 2021

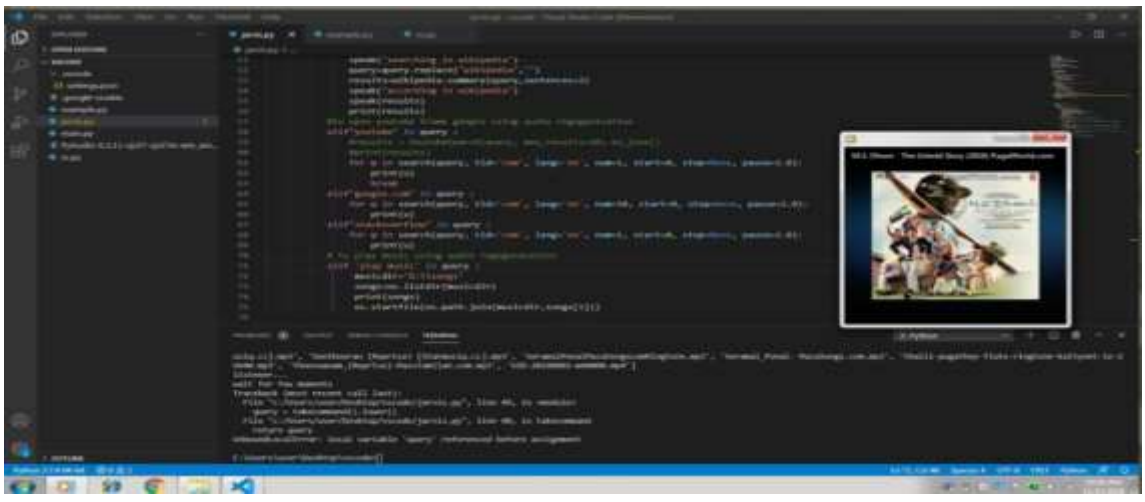
DOI: 10.17148/IJARCCE.2021.101015

Step4: **According to user command output of Google**



Step5: **According to user command output of youtube**

Step6: **Fetch the internal storage like songs,documents,files etc.,**



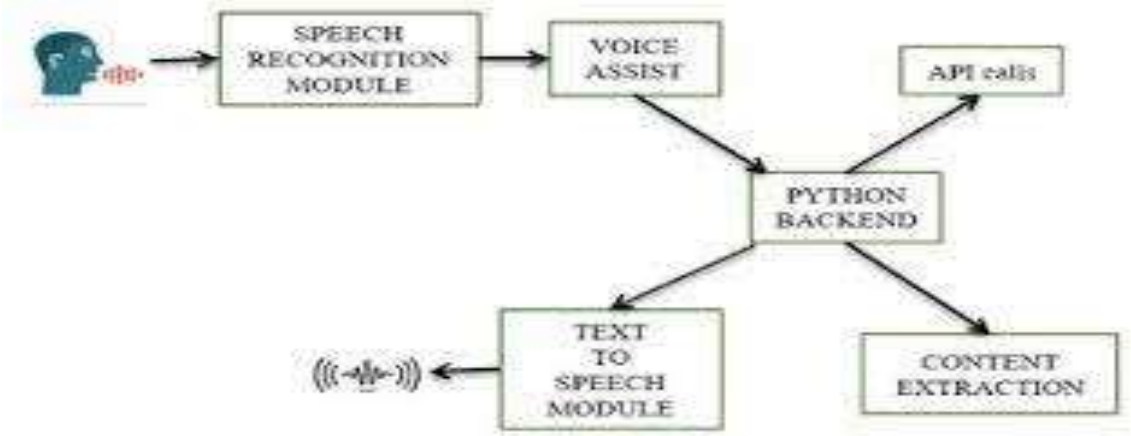
IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Vol. 10, Issue 10, October 2021

DOI: 10.17148/IJARCCE.2021.101015

Work Flow diagram:



CONCLUSION:

In this paper we have discussed a Voice Activated Personal Assistant developed 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.

FUTURE ENHANCEMENTS:

This work can be taken into more detail and more work can be done on the project in order to bring modifications and additional features. The current software doesn't support a large vocabulary. The work will be done in order to accumulate more number of samples supports only few areas of the notepad but more areas can be covered and effort will be made in this regard.

REFERENCES:

[1] M. Bapat, H. Gune, and P. Bhattacharyya, "A paradigm-based finite state morphological analyzer for marathi," in Proceedings of the 1st Workshop on South and Southeast Asian Natural Language Processing (WSSANLP), pp. 26–34, 2010.

[2] B. S. Atal and L. R. Rabiner, "A pattern recognition approach to voiced unvoiced-silence classification with applications to speech recognition," Acoustics, Speech and Signal Processing, IEEE Transactions on, vol. 24, no. 3, pp. 201–212, 1976.

[3] V.Radha and C. Vimala, "A review on speech recognition challenges and approaches," doaj. org, vol. 2, no. 1, pp. 1–7, 2012. [4] T. Schultz and A. Waibel, "Languageindependent and language adaptive acoustic modeling for speech recognition", Speech Communication, vol. 35, no. 1, pp. 31–51, 2001.

[5] J. B. Allen, "From lord rayleigh to shannon: How do humans decode speech," in International Conference on Acoustics, Speech and Signal Processing, 2002.

[6] M. Bapat, H. Gune, and P. Bhattacharyya, "A paradigm-based finite state morphological analyzer for marathi," in Proceedings of the 1st Workshop on South and Southeast Asian Natural Language Processing (WSSANLP), pp. 26–34, 2010.

[7] G. Muhammad, Y. Alotaibi, M. N. Huda, et al., pronunciation variation for asr: A survey of the "Automatic speech recognition for bangla digits," literature," Speech Communication, vol. 29, no. in Computers and Information Technology, 2009. 2, pp. 225–246, 1999.

[8] S. R. Eddy, "Hidden Markov models," Current opinion in structural biology, vol. 6, no. 3, pp. 361–365, 1996.

[9] excellent style manual for science writers is "Speech recognition with flat direct models," IEEE Journal of Selected Topics in Signal Processing, 2010.

[10]Srivastava S., Prakash S. (2020) Security Enhancement of IoT Based Smart Home Using Hybrid Technique. In: Bhattacharjee A., Borgohain S., Soni B., Verma G., Gao XZ. (eds) Machine Learning, Image Processing, Network Security and Data Sciences. MIND 2020. Communications in Computer and Information Science, vol 1241. Springer, Singapore. https://doi.org/10.1007/978-981-15-6318-8_44

[11]S. Srivastava and S. Prakash, "An Analysis of Various IoT Security Techniques: A Review," 2020 8th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO), 2020, pp. 355- 362, doi: 10.1109/ICRITO48877.2020.9198027 [12]Saijshree Srivastava, Surya Vikram Singh, RudrendraBahadur Singh, Himanshu Kumar Shukla," Digital Transformation of Healthcare: A blockchain study" International Journal of Innovative Science, Engineering & Technology, Vol. 8 Issue 5, May 202

