Voice Assistant

Ms. Aparna A.Patil[1], Shreya Shrimandhar Demapure[2], Shweta Deepak Bhosale[3]
[1]Assistant Professor, Dept of Information Technology
[2]Dept of Information Technology
[1, 2] Sharad Institute of Technology Polytechnic, Yadrav-Ichalkaranji.

Abstract- *The purpose of this study is to discuss the development of Python-based Intelligent Software Assistant application for visually challenged or blind people and for everyone. The application is intended to help people with visual limitations or blind people to access Android-based devices so that they can use library resources by using android devices. Design/methodology/approach the necessary data are collected from journals, articles, books and questionnaires, and similar applications are analyzed. The application design method used is the Scrum method, which consists of Backlog, Sprint and Scrum Meeting. From the operational side of the application, the method used is speech-to-text and text-to-speech. Findings This application has been tried with some users who have total blindness and low vision, and all provided a good response to this application. From the performance side, the user gives a very satisfied response to this application. While the ease of using the application, the user also provides a satisfactory response to the ease of using this application. Research limitations/implicationsthe application still has limitations in penetration to the user, and the application is only built using Python as its platform. In addition, the dependence on libraries from Google has caused difficulties in implementing this application with local dialect, which is only understood by the local community. Social implication This application has implications for the society, especially those with limitations in eyesight can be so much more productive and independent. This can reduce the social burden in society. Originality/value This application provides an easy access of an android device to blind people and people with low vision, as well as others to access to library resources with devices that have been installed with this application. This facility can improve the library accessibility to the blind and visually challenged community and also the normal people. Artificial intelligence technologies are beginning to be actively used in human life, this is facilitated by the appearance and wide dissemination of the Internet of Things (IOT). Autonomous devices are becoming smarter in their way to interact with both a human and themselves. New capacities lead to creation of various systems for integration of smart things into Social Networks of the Internet of Things. One of the relevant trends in artificial intelligence is the technology of recognizing the natural language of a human. New insights in this topic can lead to new means of natural human-machine interaction, in which the machine would learn how to understand human's language, adjusting and interacting in it. One of such tools is voice assistant, which can be integrated into many other intelligent systems.*

Keywords- Python, Visually Blind, voice assistance, Speech-to-text, text-to-speech.

I. INTRODUCTION

The rapid evolution of AI and machine learning made possible the development of voice recognition technology, which is how actively penetrating every area of our lives. And there's nothing to wonder about: for such social-dependent creatures as humans, speaking is a lot more natural activity than writing or, of course, typing. An average human can type about 40 words in one minute, but pronounce 150. This contrast, alongside with many other benefits we're discussing below, vividly demonstrates why voice technology should be taken seriously by not only IT giants and fans of smart homes, but every business owner out there.

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 world, we train our machines to do their tasks by themselves or to think like humans 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 their users to control their machine by just giving input in the form of 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 basic tasks that are supported by most of the virtual assistants are:

- Checking weather updates
- Sending and checking mails

- Search on Wikipedia
- Make and receive calls
- Stream music
- Open applications
- Text messages etc.

What is a Voice Assistant?

Voice assistant (might be also called digital, virtual or AI assistant) is a task-oriented programming application that recognizes human speech and carries out commands pronounced by a user. It is powered by AI and bases its performance on cloud storage with millions of words and phrases in it. Unlike the very first voice recognition devices that scientists were working on in the 1940-50s, modern digital assistants aren't restricted by a certain language pattern or vocabulary. As in 2018, there are two types of voice assistant software:

- Consumer-facing voice assistants for homes;
- Chatbots and virtual mobile assistants for a workplace.

The first type belongs to home automation (also called domotics or smart houses), a system that allows controlling appliance, lighting, electricity, gadgets, and other things filling in our houses. Home automation requires an internet connection and is a part of the Internet of Things (IoT). The second type, voice assistants for a workplace, can be used out of touch with building automation, as a standalone application incorporated in the professional routine of a team.

Aim

In the previous days people done all work manually, suppose because of any reason people is not able to done work manually then using voice assistant they can easily access any information or done any work. So, for that reason we will implement this system.
-Also used for blind people.

Scope

- To done our all-daily programs more easily.
- Using voice, we can perform many tasks so that is very useful for blind people to operate apps easily.

II. METHODOLOGY

By developing this app, we can overcome the problems like the will able to get the date and time, they can also search the information on google, they can visit YouTube, and they can play the music whenever they need, they can also create their own notes, they can also create the mails and send to people.

If you have ever wondered how to create a virtual assistant software, the foremost thing to do would be to get familiar with how ASR works. In a nutshell, the process starts with the device gathering audio with the microphone. Recorded speech waveforms get straight to acoustic analysis, which is performed on three different levels:

- acoustic modeling, which represents the which phonemes were pronounced and what are the words these phonemes complete;
- pronunciation modeling, that analyzes the way phonemes are pronounced, is there any accent or other peculiarities of the vocal apparatus to capture the phonetic variability of speech;
- language modeling, which is aimed at finding contextual probabilities depending on what phonemes were captured.

III. BLOCK DIAGRAM

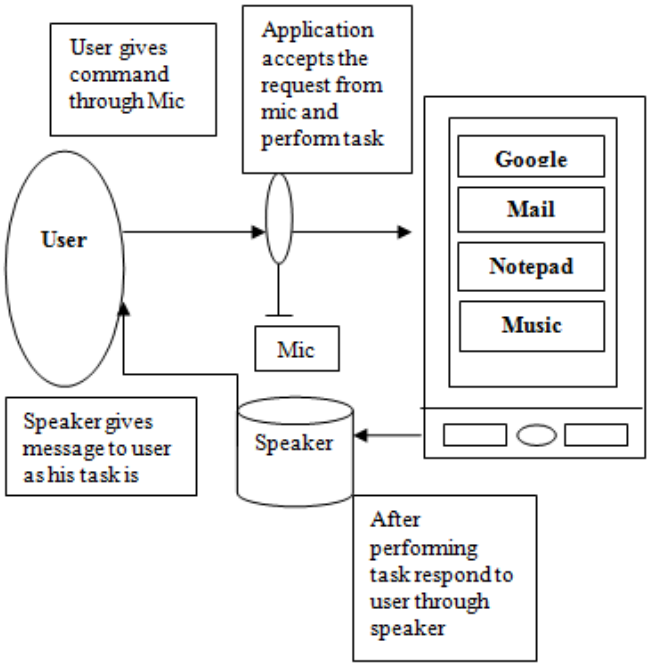

Fig- Block diagram of Voice assistance

IV. PROBLEM DEFINITION

As blind person has to depend on others for their day-to-day life so often in our society, that people consider them as a burden. And often it happens that blind people also want to

live their life as a normal people such as using mobile's and laptopsetc. to perform some activities so they need any device that can help them to perform the activities.

V. LITERATURE 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, Bluetooth 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.

1. Jarvis Technique:

Jarvis is an AI system Zuckerberg has built to control his home and perform basic tasks, such as turning the lights off or on, control a particular room's temperature, playing music, opening doors, and so on. The home AI uses natural language processing and speech recognition to understand the user's voice, and the context of the command, in order to perform the tasks, it is asked to do. It also has facial recognition capabilities and, therefore, can open doors of the home for known guests, or tell Zuckerberg where each member of the family is, whether Max is asleep or playing around, and so on.

2. Conversion of speech to text:

When user gives command to application, after detecting that command it should convert that command i.e., speech into the text format. And then it performs task. after completing the task, it should convert that command into text.

3. Conversion of text to speech:

As blind person has to depend on others for their day-to-day life so often in our society, that people consider them as a burden. And often it happens that blind people also want to live their life as a normal people such as using mobiles and laptops etc. to perform some activities so they need any device that can help them to perform the activities.

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, Bluetooth 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. 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.

Resources Required:

a) Hardware Resources

- A computer system with installed Voice Assistant application
- Earphones.
- Mic

b) Software Resources

- Visual Studio
- Database
- Python IDL

Algorithm:

- **Give Commands to device:** First the user gives command to the devise which is in the form of speech.
- **Listen Commands:** After giving the command by the user the device will listen to the commands of the user carefully.
- **Recognizes speech/command:** After listening the command it will first recognize it.
- **Conversion of Speech-to-text:**As it recognizes the speech it will convert the given speech into the text format.
- **Perform task:** After conversion from speech to text it performs the required operations on the text to develop/give the desired output to the user.
- **Conversion of text to speech:** After doing all the operations it converts the text into the speech form and delivers it to the user as the output is required.

VI. PROPOSED SYSTEM

The proposed system will have the following functionality:

(a) The system will keep listening for commands and the time for listening is variable which can be changed according to user requirements.
(b) 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 can have both male and female voices according to user requirements.
(d) Features supported in the current version include playing music, emails, texts, search on Wikipedia, or opening system installed applications, opening anything on the web browser, etc.
(e)The system will keep listening for commands and the time for listening is variable which can be changed according to user requirements.
(f) 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.

VII. FUTURE SCOPE

The virtual assistants which are currently available are fast and responsive but we still have to go a long way. The understanding and reliability of the current systems need to be improved a lot. The assistants available nowadays are still not reliable in critical scenarios. The future of these assistants will have the virtual assistants incorporated with Artificial Intelligence which includes Machine Learning, Neural Networks, etc. and IoT. With the incorporation of these technologies, we will be able to achieve new heights. What the virtual assistants can achieve is much beyond what we have achieved till now. Most of us have seen Jarvis, that is a virtual assistant developed by iron man which is although fictional but this has set new standards of what we can achieve using voice-activated virtual assistants.

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

REFERENCES

[1] DO'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] "Speech recognition with flat direct models," IEEE Journal of Selected Topics in Signal Processing, 2010.
[3] Amrita S.Tulshan, Sudhir NamdeoraoDhage, "Survey on Virtual Assistant: Google Assistant, Siri, Cortana, Alexa".
[4] Gaurav Agrawal, Harsh Gupta, Divyanshu Jain, ChinmayJain, Prof. Ronak Jain, irjmets, "Desktop voice Assistant".