Desktop Voice Assistant for Visually Impaired

Ankush Yadav, Aman Singh, Aniket Sharma, Ankur Sindhu, Umang Rastogi

Abstract: 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).

I. INTRODUCTION

The usage of virtual assistants is expanding rapidly after 2017, more and more products are coming into the market. Due to advancement in the technology many different features are being added in the mobile phone and desktops. To use them with more convenient and fun way we require a means of input which is faster and reliable at the same time. In our project we use voice command to input the data into the system for that the microphone is used which convert acoustic energy into electrical energy. After taking the input there is a requirement to understand the audio signal for this google API is used. Different companies like google, apple use different API's for this purpose. It is truly a feat that today, one can schedule meetings or send email merely through spoken commands.

Ankush Yadav, Department of Computer Science and Engineering Meerut Institute of Engineering and Technology, Baghpat Bypass, Meerut, U.P

Aman Singh, Department of Computer Science and Engineering Meerut Institute of Engineering and Technology, Baghpat Bypass, Meerut, U.P

Aniket Sharma, Department of Computer Science and Engineering Meerut Institute of Engineering and Technology, Baghpat Bypass, Meerut, U.P

Ankur Sindhu, Department of Computer Science and Engineering Meerut Institute of Engineering and Technology, Baghpat Bypass, Meerut, U.P

Umang Rastogi, Department of Computer Science and Engineering Meerut Institute of Engineering and Technology, Baghpat Bypass, Meerut, U.P

These voice assistants work as your companion which performs your day by day task with minimum efforts and also help the user to function better by giving daily updates. It was after the recognition of importance of voice commands in day to day life that we have aimed to develop a personal assistant for desktop which will do every work from playing music to sending an Email.

II. RELATED WORK

The development of voice assistants was started in 1962 at the Seattle world's fair where IBM presented a device called shoebox IBM that could recognize spoken digits and then return them back through igniting lamps labelled next to the digits 0-9. It had the ability to perceive a total of 16 words. Currently most of the voice assistances are developed for the mobile phones like google make voice assistance support for android mobiles, apple use Siri and amazon have Alexa these assistants used language processing to perform its task. another voice assistant is Cortana which is been developed by Microsoft and used on desktop. All of these voice assistants perform the same intended function – that is, voice initialized processing, and all of these developments have been a result of the same new age technology- Artificial intelligence. At the core of all these assistants is a simple synchronous cycle – Voice commands and hear responses. Sutar Shekhar, and various researchers have jointly come up with an application which implements most of the system functionalities through voice and they also included the feature of sending a message with their voice command to help those people who are visually impaired. They aim to continuously develop their application so as to m a finally have an engine which can also recognise different local languages like Bengali, and a number of dialects of prevalent Hindi. Miss. Priyanka V. Mhamunkar and others proposed a system which will let the individual fetch meanings of words through synthesised sounds. Omyonga Kevin and Kasamani Bernard Shibwabo claim to have come up with an application that is able to implement spoken commands even without an internet connection thus, giving us flexibility over data costs.. As there is no need of internet connection this feature makes their solution faster than many of the high-profile engines like Alexa and Cortana. Tong Lai Yu and Santhrushna Gande have developed a solution for Android by using open source services, that can help programmers who are facing physical challenges in development. They used producer-consumer paradigm at the client side to integrate tasks that can be operated over an android device. On analysing these systems, we concluded that they are basically designed to work on a android platform, that's why we decide to develop a software which will work on the desktop.

We used programming language python because it is one of the most robust programming languages and with the help of pyttsx3 and speech recognition API's the development of software become easier and work with better accuracy. To help the visually impaired people our software always repeats the command which the user gives to the system so that the users are aware whether they have inserted the correct command or not. On the other hand, it also keeps on listening and fulfil the demand of the user until the user decides to quit.

III. PROPOSED WORK

i) Methodology

1. Speech recognition

The proposed system used the google API to convert input speech into text. The speech is given as an input to google cloud for processing, As an output, the system then receives the resulting text.

2. Backend work

At backend the python gets the output from speech recognition and after that it identifies whether the command is a system command or a browser command. The output is send back to python backend to give desired output to user.

3. Text to speech

Text to speech, or TTS, is a new wave technique of for transforming voice commands into readable text. Not to mix it up with VR Systems that instead, generate speech by joining strings gathered in an exhaustive DB of pre-installed text and have been developed for different goals which form full-fledged sentences, clauses or meaningful phrases through a dialect's graphemes and phonemes. Such systems have their limits as they can only determine text on the basis of pre-determined text in their databases TTS systems, on the other hand, are practically to "read" strings of characters and dole out resulting sentences, clauses and phrases.

ii) 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 stringified data is manipulated through Python Script to finalise 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.
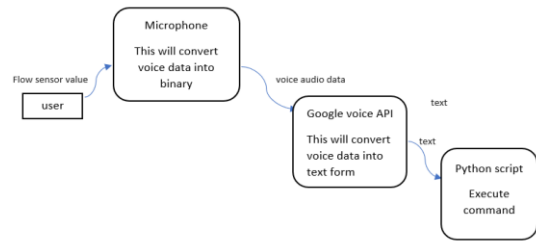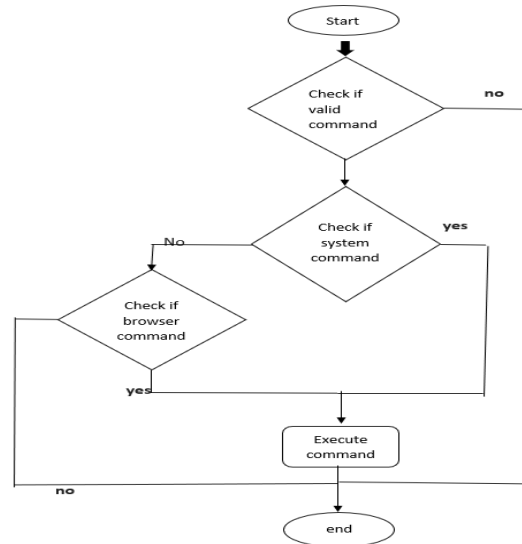

Figure1. Data flow diagram

Figure2.Flow chart of desktop voice assistant

iii) 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 pre-determined 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.

IV. VOICE ASSISTANT PRIVACY CONCERNS

A user may have a privacy concern as personal assistant require a huge amount of data and are always listening to take the command This passive data is then, retained and sifted through humans that are employed by almost all of the major companies- Amazon, Apple etc. In addition to this discovery over the Ai being able to record our audio interactions, there have been concerns over the type of data such employees and contractors were hearing. So, in the cloud-based voice assistant, privacy policy must be there which will protect the user information.

V. RESULTS

▢ This system allows user to send email in which the input is taken through voice.
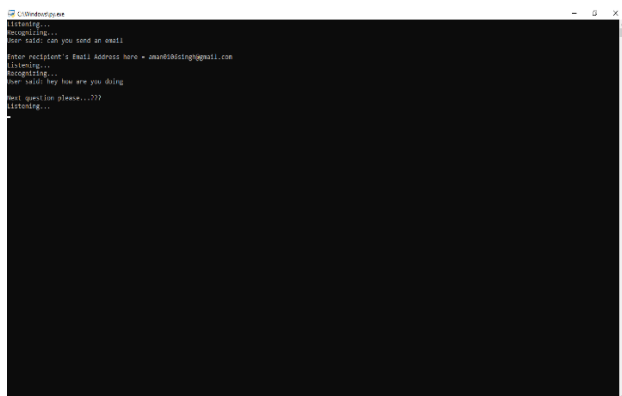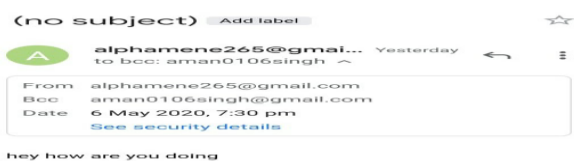

Figure3. Taking input by voice

Figure4. Email sent

▢   Users can fetch the current weather detail of that area.
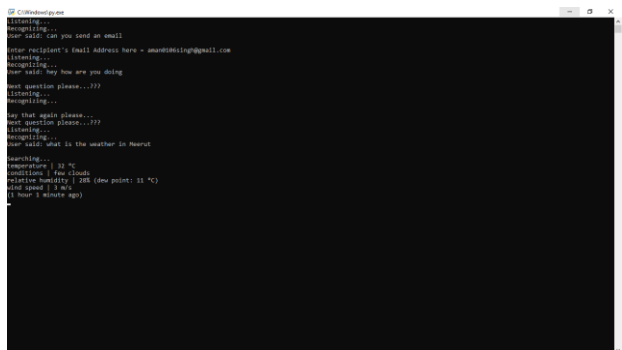

Figure5. Displaying weather condition

▢ For the entertainment purpose this system plays a random music from the folder which is specified in the program
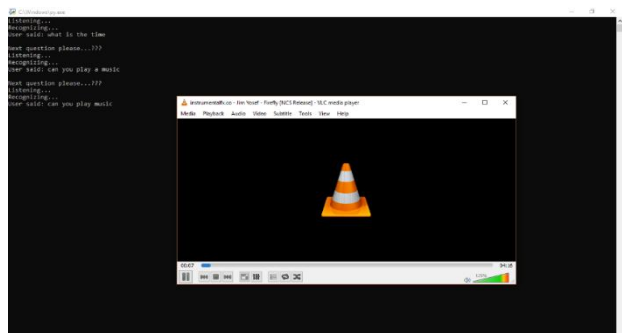

Figure6. started playing music

▢ This system allows to search anything from the internet and give the output in the form of text as well as speech
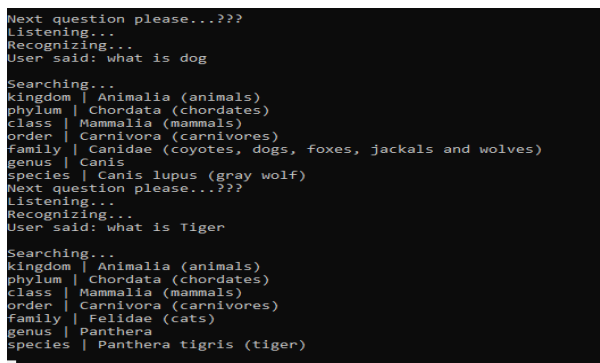

Figure7. Information gathered from google

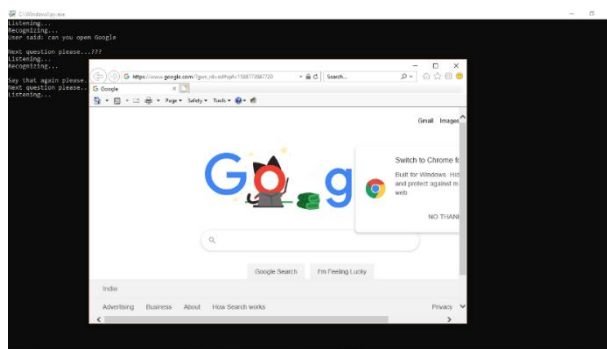▢ To facilitate the user the proposed system allow to open different applications just by using a voice command

Figure8 opened google chrome

VI. CONCLUSION

Voice assistants have had a huge change in user' s interaction with technologies embedded in their devices. Like any other technology of such magnitude, they have altered the basic genome of the sphere in which they operate. While this has largely created a better world with drastic benefits for communities, which were before kept in dark with reference to technological innovations, they have posed new kind of threats with respect to user's privacy and security.

VII. FUTURE SCOPE

The future of voice assistants can be parameterised on an array of dimensions. With respect to compatibility and integration with other devices and systems, there i sill a lot to be achieved, Another dimension would be with respect to the redundant use of wake words at the beginning of each command. The individuality of results also poses a big problems.

But for all intents and purposes, the future of these technology is a bright one. With advances in it and in technologies related to it (search processes, for example) Voice assistants can carry out even more complex tasks like booking tickets, etc.

At its core, this technology might have its own trials and tribulations, but it is still a boon for many who might have been kept in the dark with all spheres of technological developments.

Apart from this, it is just too beneficial a technology to not go through continuous research and development.

REFERENCES

1. Sutar Shekhar, Pophali Sameer, Kamad Neha, Deokate Laxman, intelligent voice assistant using Android platform. International Journal of Advance research in computer Science and Management Studies. Volume 3, Issue 3, march 2015
2. Mhamunkar, M. p. v., Bansode, M. k. S., & Naik, L.S. (2013). Android application to get word meaning through voice. International journal of Advance Research in Computer Engineering & Technology (IJARCET). 2(2), pp-572.
3. Apte, T. V., Ghosalkar, S., pandey, S., & Padhra, S. (2014). Android app for blind using speech technology. International Journal of Research in Computer and Communication Technology (IJRCCT), #(3), 391-394.
4. Anwani, R., Santuramani, U., Raina, D., & RL, P. Vmail: voice Based Email Application. International Journal of Computer Science and Information Technologies, Vol. 6(3), 2015
5. Deepak shende, Ria Umahiya, Monika Raghorte, Aishwarya Bhisikar, Anup Bhange.AI based voice assistant using python. JETIR, feburary 2019, vol 6, issue 2.
6. Thakur, N., Hiwrale, A., Selote, S., Shinde, A. and Mahakalkar, N., Artificially Intelligent Chatbot
7. Yu, T. L., Gande, S., & Yu, R. (2015, January). An open -Source Based Speech Recognition Android Application for Helping Handicapped Students Writing Programs. In Proceedings of the International Conference on Wireless Networks (ICWN) (p. 71). The Steering Committee of The World Congress in Computer Science, Computer Engineering And Applied Computing (WorldComp).
8. Yannawar, P. (2010). Santosh K. Gaikwad Bharti W. Gawali Pravin Yannawar. A Review on Speech Recognition Technique. International journal of Computer Application
9. Brandon Ballinger, Cyril Allauzen, Alexander Gruenstein, Johan Schalkwyk, On-Demand Language Model Interpolation for Mobile Speech Input INTERSPEECH 2010, 26-30 September 2010, Makuhari, Chiba, Japan, pp 1812-1815.
10. IOSR Journal of Engineering Mar. 2012, Vol. 2(3) pp: 420-423 ISSN: 2250-3021 www.iosrjen.org 420 | "Android Speech to Text Converter for SMS Application" Ms. Anuja Jadhav* Prof. Arvind Patil**.
11. Michael Stinson, Sandy Eisenberg, Christy Horn, Judy Larson, Harry Levitt, and Ross Stuckless" REAL-TIME SPEECH-TO-TEXT SERVICES."
12. International Journal of Information and Communication Engineering 6:1 2010 "The Main Principles of Text-to-Speech Synthesis System",K.R. Aida–Zade, C. Ardil and A.M. Sharifova.
13. Review of text-to-speech conversion for English, Dennis H. Klatt, Room3 6-523, Massachuset Institute of Technology Cambridge Massachusetts.
14. 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.
15. Shibwabo, B. K., & Omyonga, K. (2015). The application of real-time voice recognition to control critical mobile device operations.

AUTHOR PROFILE

Umang rastogi at present working as a assistant professor in computer science and engineering department in MIET engineering college, Meerut, UP, India. He is having more than 11 years' experience in teaching field at national and international universities/colleges' have published multiple research papers in Scopus and peer reviewed journals. His research area domain is computer networks and AI/ML. He earned bachelors and master's degree in computer science and engineering department with first division. His master's thesis work in routing protocols. Where he did deep investigation on performance of DSR protocol over MANET with correlation method. He worked in multiple tools like network simulators (NS2 and NS3) and IBM Watson studio.

Ankush Yadav a student of the 4th year from Meerut Institute of Engineering and Technology, Meerut under the program of Bachelor of Technology in Computer Science and Engineering. He has worked and have interest in technologies like Python, Core Java, C programming. He is core java certified from NIIT. He has also done a Ethical hacking training from Internshala. He is fun loving person who is always ready to learn something new.

Aman Singh a student of the 4th year from Meerut Institute of Engineering and Technology, Meerut under the program of Bachelor of Technology in Computer Science and Engineering. He has worked and have interest in technologies like Python, Core Java, C programming.. He has also done a Python training from UDEMY. He is always eager to learn new things about the new technology.

Aniket Sharma a student of 4th year from Meerut Institute of Engineering and Technology, Meerut under the program of Bachelor of Technology in Computer Science and Engineering. He has worked and has interest in technologies like Python, Core Java, C programming and Angular. He has also done a Python training from Private Institute. He is currently working as an Angular software developer at Appinventiv technologies Pvt. ltd.

Ankur Sindhu a student of the 4th year from Meerut Institute of Engineering and Technology, Meerut under the program of Bachelor of Technology in Computer Science and Engineering. He has worked and have interest in technologies like Python, Core Java, C programming..