Desktop Based Voice Assistant Application Using Machine Learning Approach

Ashish Agarwal[1], Chanchal Sharma[2], Tushita Agarwal[3]

[1,2,3] *Student, Dept. of Computer Science Engineering, MIET, Meerut, UP, India*

---***---

Abstract - *A voice assistant can schedule reminder , it can note things , it can calculate , it can search and do more stuffs by Desktop Voice assistant system. This project works on input - output system it takes voice as input and gives output in form of voice or by displaying text on screen. This system makes human being intelligent and provide us an quick result Also saves the time . Microphone is used to take inputs and then that input is converted into a computer language and give the best result.*

Key Words: **Desktop voice assistant, Natural language processing, Microphone, Computer language, Machine learning.**

1. INTRODUCTION

Ai system are more popular in today's world as it help in interaction of human being with machines. Machines can also now interact , observe , explore actions ,habits this shows how much development is there. The software programs that help us to ease our day to day task is virtual assistant. Command can be given in two forms by text or by voice. Voice based smart assistant need and invoking word. Examples of these types of assistant are Apple's siri , Amazon's alexa , Microsoft's cortana etc[1]. This project is based on the desktop. In this generation where technology is overpowering , machines has the major role. Performance is the main factor. Technology is developing in such a way that it can behave like human and do work by own. This is how virtual assistant came into picture[2]. Digital assistant that use speech recognition to take commands of user and by the language processing algorithm process the commands and perform the task as per user requirement. Virtual assistant ignore the ambient noise and gives the relevant information[3,4]. This is a complete software based technology but it is also being integrated into the devices and also can work on single device such as alexa. Major changes are occuring in technology therefore it is necessary to train in machine learning , deep learning and neural networks. Today we interact with our machine via voice assistant. Nowadays , all big companies are using virtual assistant so that they can use voice to ease their work. Hence , by voice assistant we are advancing to higher level through which we can interact with machine[5,6]. This technology is more useful for the old people , children , physically challenged and blind people as it make it ease to interact with machines

and voice is the major key . It is most useful for the blind people who can just use their voice to do things[7]. Examples of the task we can do with virtual assistant :- • Weather updates • Update about mail • Web search • Can play music, video • Alarm setter • Run any program or application • Reading newspaper Not just these task we can do much more as per the need of the user. In this project the voice assistant is developed for the window or linux users. It is a desktop-based assistant that is developed using python libraries and modules . This assistant can do all basic tasks as it is the basic model , current technology is good but need to be merged with IOT and machine learning for advance feature and enhancements[8]. Python libraries and modules are used to develop the model and machine learning is used to train the model , linux and windows commands are also being used in this so that it can work perfectly on the operating system. Three models that involved in our model is:- 1. Supervised Learning 2. Unsupervised Learning 3. Reinforcement Learning These model are depending on the usage of the user. Machine learning and deep learning play important role in achieving this. Writing commands again and again for particular task is completely removed by the voice assistant. As the model is created any number of users can use it at any time and easy to use . Voice assistant help to do any task at the ease and also with less time[9].

2. LITERATURE REVIEW

Voice assistants had come into the picture from 1962. IBM brought a tool into the market known as shoebox IBM . This device can recognise the spoken digits , which it return through the lamps. It can grasp 16 words in total[10]. All assistants use language processing to execute its tasks. Further we have a voice assistant named cortana , it had been developed by Microsoft for the desktop purpose. So the basic function of all the voice assistants is to voice configure processing , and this development is the result of the modernistic technology which is AI(artificial intelligence) . Crux of these assistant is a coincident cycle i.e voice command. A well known personality Sutar Shekhar and various researchers together bring an app which implies the system functioning by voices[11]. It further involve the ability which is send message through voice command , it is specially for the people who are partially sighted . In future there is a plan to develop an engine who recognise local language. As we all know that

python language is one of the powerful and easy language therefore we used python language. Speech recognition and the Pyttsx3 module is used in developing s\w becomes more trouble -free and work in more precise way[13]. For the partially sighted people our software has one feature which is really helpful for them that is repeating the commands given by the user so that the user know that they have given the right commands. Also our software continues listening and processing demands of users till the end[12].

3. PROPOSED METHODOLOGY

We can change the voices of voice assistants by our choice i.e. male or female. It keeps listening to the commands and we can change the time of listening by yourself i.e. it is variable. Speech Recognition - It converts input speech into text through google API. Speech input will be given to google cloud for processing from which text will be received by system. Backend Work - The output which is generated by speech recognition given to python at backend. Then it recognise it as a system or browser instruction. On solving it gives the required output to the user. Text Data Into Speech Data - It is a technique which is used to transform voice commands into readable text. It is a novel way technique. It is not similar to VR systems. System like this have its own limitations. It converts text data into the words by processing the NLP. The below data flow diagram explains the path. Users give the voice command through the microphone which converts voice data into the binary data and then through the speech recognition module it converts voice data into the text and at last python executes the command.
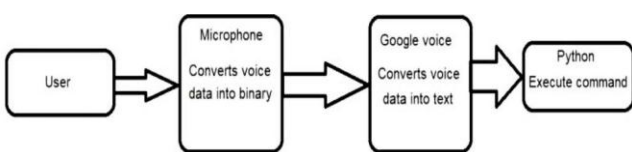

Figure1: Data Flow Diagram

Voice assistants take speech patterns as the input from the microphone. Then it recognises the audio and converts the audio into the text format. Comparison is done of inputs. And at last it gives the expected output. In starting data is taken as speech and in result it gives the text data by using the NLP. Further , resultant string data is operated by Python to conclude the final result. Lastly, output that is processed will be present in text & in form of speech by conversion through TTS. The below voice flow diagram explains the flow of the input command. The voice command is reorganized by the speech recognition module which calls the backend of the python in which some APIs are called and some modules are also called like text to speech module and content extraction.
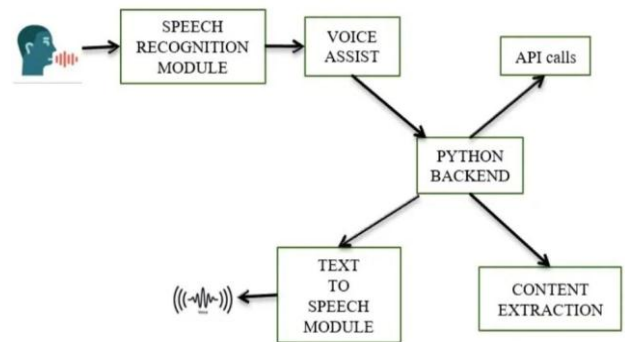

Figure 2: Voice flow diagram

The below Flow chart of the voice assistant explains the path of the commands. If the command is valid then, it checks that the given command is a system command or not. If that command is not the system command then it checks it as a browser command. If it is a browser command then it will execute the command otherwise it goes for the new command.
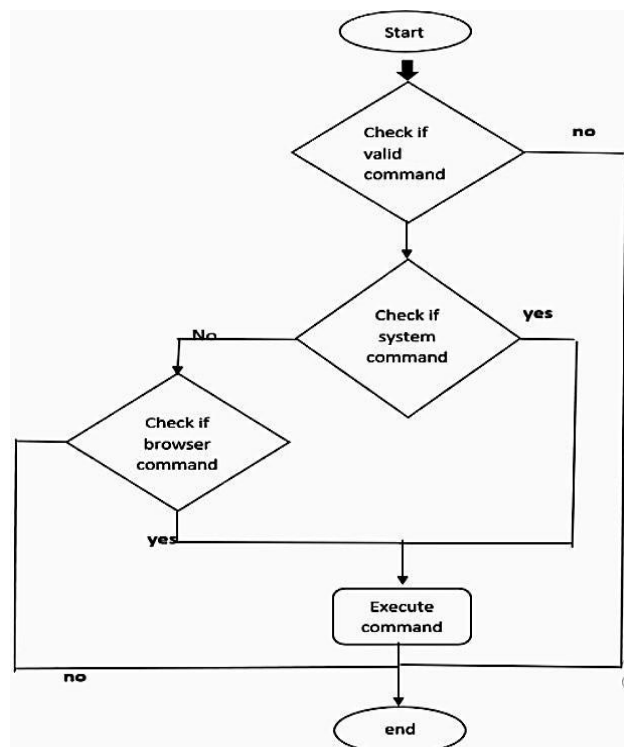

Figure 3 : Flow Chart of Voice Assistant

4.RESULT

It takes voice input from the user and then evaluates it and at last it gives the desired output. Below, there are some user inputs in figure 4 and figure 5.

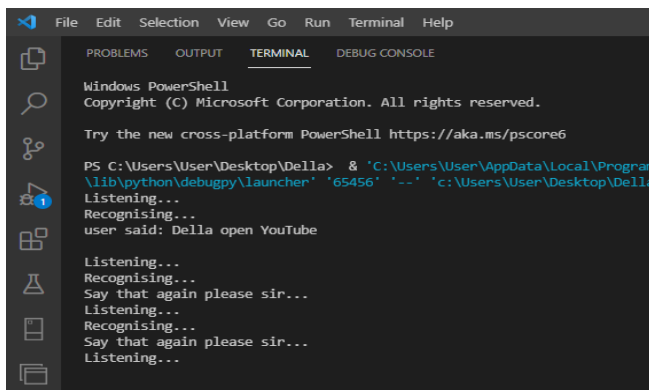● This system allows users to open youtube using voice command.
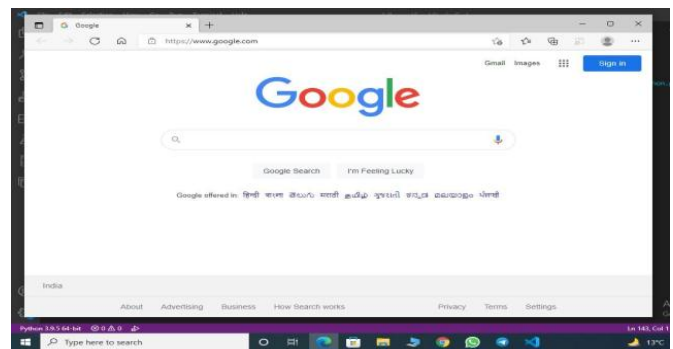
Figure 4: It takes command to open the youtube

● This system allows users to open Google by voice command.
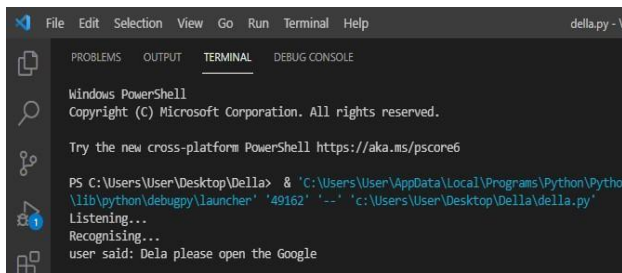
Figure 5: It takes command to open the Google

Now, it gives the desired outputs of above mentioned commands. In figure 6 and figure 7 it shows the expected output.
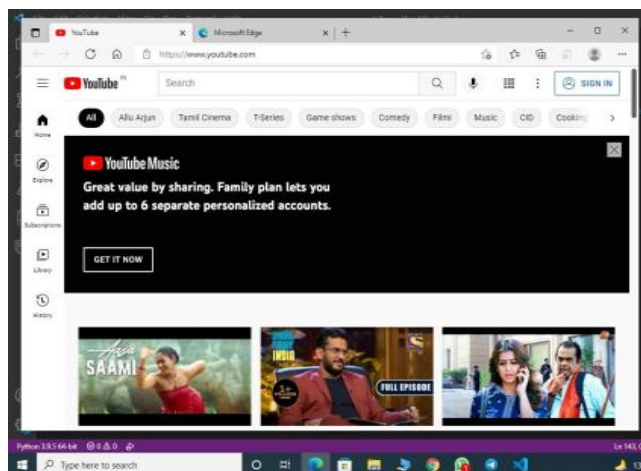
● So, it opened Youtube.

Figure 6: It gives the output of the above command

● So, it opened Google.

Figure 7: It gives the output of the above command

● It also searches for information from wikipedia.

Figure 8: It gives the output through the wikipedia

5.CONCLUSION & FUTURE SCOPE

In this research paper we explored about the design and implementation of the digital assistance. In this project the open source libraries and python is used. The more features can be added to this project without disturbing the present system. This project reduces the manual work. It just not used for human commands but also takes user queries like opening and other operations. Voice assistants have a very bright future. There is a lot of work on voice assistants to be achieved. Voice assistants can carry out even more difficult tasks like taking appointments through voice commands, booking tickets, playing or stopping the audio/video. Voice assistants are the type of remote in our lives.

6.REFERENCES

[1] R. Belvin, R. Burns, and C. Hein, "Development of the HRL route navigation dialogue system," in Proceedings of ACL-HLT, 2001

[2] V. Zue, S. Seneff, J. R. Glass, J. Polifroni, C. Pao, T.J.Hazen,and L.Hetherington, "JUPITER: A Telephone Based Conversational Interface for Weather Information," IEEE Transactions on Speech and Audio Processing, vol. 8, no. 1, pp. 85–96, 2000.

[3] M. Kolss, D. Bernreuther, M. Paulik, S. St¨ucker, S. Vogel, and A. Waibel, "Open Domain Speech Recognition & Translation: Lectures and Speeches," in Proceedings of ICASSP, 200

[4] 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

[5] 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.

[6] 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.

[7] 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

[8] Deepak shende, Ria Umahiya, Monika Raghorte, Aishwarya Bhisikar, Anup Bhange.AI based voice assistant using python. JETIR, february 2019, vol 6, issue 2.

[9] Thakur, N., Hiwrale, A., Selote, S., Shinde, A. and Mahakalkar, N., Artificially Intelligent Chatbot

[10] 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).

[11] Yannawar, P. (2010). Santosh K. Gaikwad Bharti W. Gawali Pravin Yannawar. A Review on Speech Recognition Technique. International journal of Computer Application

[12] 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.

[13] 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**.

[14] Michael Stinson, Sandy Eisenberg, Christy Horn, Judy Larson, Harry Levitt, and Ross Stuckless" REAL-TIME SPEECH-TO TEXT SERVICES."

[15] 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.

[16] Review of text-to-speech conversion for English, Dennis H. Klatt, Room 3 6-523, Massachusett Institute of Technology Cambridge Massachusetts.

[17] 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.

[18] Shibwabo, B. K., & Omyonga, K. (2015). The application of real time voice recognition to control critical mobile device operation

BIOGRAPHIES

Ashish Agarwal is a final year graduate in Computer Science and Engineering at Meerut Institute of Engineering and Technology, Meerut, U.P, India.

Chanchal Sharma is a final year graduate in Computer Science and Engineering at Meerut Institute of Engineering and Technology, Meerut, U.P, India.

Tushita Agarwal is a final year graduate in Computer Science and Engineering at Meerut Institute of Engineering and Technology, Meerut, U.P, India.