POODLE: Voice Assistant Using Raspberry Pi

Dhrati Gupta[1], Abhishek Dixit[2], Akshat Mitra[3], Mr. Anuj Kumar[4]
[1,2,3]*UG Final Year Students, Department of Electronics & Communication Engineering, RKGIT College, U.P., India*
[4]*Assistant Professor, Department of Electronics & Communication Engineering, RKGIT College, U.P., India*

Abstract **- Voice Assistant and Smart Speakers bring a replacement dynamic into the house. Their tongue capability and straightforward use simplify tasks we perform each day, from booking a gathering, playing music or making an internet purchase. Their convenience keeps us heads-up and hands-free. A voice assistant could also be a digital assistant that uses voice recognition, language processing algorithms, and voice synthesis to concentrate on specific voice commands and return relevant information or perform tasks as requested by the user. Apple's Siri, Amazon's Alexa, Microsoft's Cortana, and Google's Assistants are the foremost popular voice assistants and are embedded in smartphones or dedicated home speakers. Users can ask assistants questions, perform home automation and media playback using voice, and manage other basic tasks such as email, to-do lists, and calendars via voice commands. we will make a fully functional one employing a Raspberry Pi.**

Index Terms **- Personal Assistant, Text to speech, Speech to text, Raspberry Pi, Voice Command System, Query Processing Machine Learning.**

INTRODUCTION

For most folks, the last word of luxury would be an assistant who always listens for your call, anticipates your every need, and takes action when necessary. That luxury is now available in the form of AI assistants, aka voice assistants. Voice assistants are available in somewhat small packages and may perform a spread of actions after hearing a wake word or command. They can turn on and turn off lights, interact with users, play music, place online orders, etc. Voice assistants aren't to be confused with virtual assistants, which are people that work remotely and may therefore handle all types of tasks. Rather, voice assistants are technology-based. As voice assistants become more robust, their utility in both the private and business realms will grow also.

Apple's Siri has been around the longest, released as a standalone assistant in 2010 and bundled into iOS in 2011. Microsoft followed shortly thereafter by launching Cortana in 2013. Amazon launched Alexa with its Echo-connected home speaker in 2014, and Google's Assistant was announced in 2016 alongside its Home speaker and is additionally embedded within the Google app for Android-based smartphones. Each assistant has its own unique features, but the core functions are equivalent. Voice assistants differ from earlier voice-activated technologies therein they will answer a larger number of commands and questions. This is because they are always connected to the Internet; each interaction is shipped back to a central computer system that analyzes the user's voice commands and provides the assistant with the proper response. Earlier voice assistant devices relied on a smaller set of "built-in" commands and responses. Recent advances in tongue processing, also referred to as linguistics, have allowed voice assistants to make meaningful responses quickly. Hirschberg and Manning credit the recent improvements in NLP to four things: (i) a vast increase in computing power, (ii) the availability of very large amounts of linguistic data, (iii) the event of highly successful machine learning (ML) methods, and (iv) a way richer understanding of the structure of human language and its deployment in social contexts.

Voice assistants are available on a good sort of hardware platform. Amazon and Google both market dedicated home speaker devices for his or her voice assistants. Amazon makes several variations of its Echo product, from the small Echo Dot to the Echo Show, which has audio and video capabilities. Google's speaker hardware is named the house and

also comes in mini- and full-size models. Apple is simply entering the house speaker market, with the announcement of its Siri-enabled HomePod device, scheduled to be available in December 2017. Microsoft has built Cortana into Windows 10 PCs and phones and recently partnered with Harmon Kardon to develop a Cortana-enabled home speaker. Assistants are now available on most platforms as well; Google's assistant is integrated into Android phones and can be installed as a separate app on the iPhone, although some features are disabled. Amazon's Alexa has Android and iOS versions, and Microsoft and Amazon are working effectively to bring Cortana to Amazon devices and Alexa to PCs. Apple has kept their assistant off of non-iOS devices, but Siri is out there on all Apple devices, including iPhones, Macbooks, iPads, and therefore the Apple Watch. As the voice assistant market stabilizes, it seems likely that there'll be additional integration which features sets across the most voice assistants will become similar. For an instant, Amazon is the dominant player within the field, thanks to launching a home product first with an outsized media library available out of the box. Google is building capacity, and therefore the addition of a home-based speaker and integration into other Google products will drive their market share up. Apple may also become 6% to 10% relative, for the system that had the word error rate of 17% to 52%.

Text to speech conversion is the process of transforming a machine-recognized text into any language which could be identified by a speaker when the text is read out loud. It is a two-step process that is divided into the front end and back end.
The first part is responsible for converting numbers and abbreviations to a written word format. This is also mentioned as the normalization of text. The second part involves the signal to be processed into a clear one.

Speech Recognition is the ability of a machine for instance a computer to understand words and sentences spoken in any language. These words or sentences are then converted to a format that would be understood by the machine. Speech recognition is implemented using vocabulary systems and a collection of keywords. A speech recognition system could also be a little Vocabulary-many user system or an outsized Vocabulary- small user system.

2. LITERATURE REVIEW

Veton Këpuska, Next-Gen Personal Assistants (Microsoft Cortana, Apple Siri, Amazon Alexa and Google Home), Published in 2018, This paper gives us knowledge about the uses of the multi-modal dialogue systems which process two or more combined user input modes, like speech, image and interface so as to style. We moreover get broad ideas about the multiple user specific interactions. This paper gives ideas about the next generation of virtual personal assistants and modifications that can be made to interact with the assistants.

Sumit kumar Sarda, VPA: Virtual Personal Assistant, this paper describes the way to build a voice assistant that reduces the utilization of input devices such as keyboard and mouse on our PCs. This paper gives ideas about using personal assistants in our day-to-day life rather than using personal computers.

Amit Rakesh Shrivastava, "Home Server and NAS using Raspberry Pi", Published in 2017, The paper aims at providing a coffee cost NAS system which is straightforward to use and configure. It also comprises added security features and Web Server capabilities. It enables you to have round the clock available storage devices which are handy as well as power saving and allows accessibility to data on and off the network.

Zhang Huili, Realization of Files Sharing between Linux and Windows supported Samba. With the development of computer networks, there are multiple operating systems in one local area network frequently. Windows and Linux are comparatively popular in these operating systems. It is necessary to realize file sharing between different operating systems in the local area network. The Samba software package acts as the bridge between Windows and Linux.

P. Milhorat, "Building the next generation of personal digital Assistants,". Voice-based digital Assistants such as Apple's Siri and Google's Now are very demanding in today's world. Yet, despite their promise of being context-aware and adapted to a user's needs and really distinct preferences, truly personal assistants are still missing. In this paper we highlight a variety of the challenges in building personalized

voice-operated technology and propose a number of R&Ds we've undertaken so as to solve them. Particularly, we concentrate on natural language and dialog management aspects as these parts of the technology pipeline require the largest amount of augmentation.

Cloud storage requires internet connectivity. Hence there is a need for a convenient storage device which can be accessed by multiple users simultaneously. In the current proposed system, we are developing better and efficient systems which can provide storage and access to data simultaneously to the end-users in various numbers and also serve the aim of voice assistants being used in our day-to-day life.

3. PROPOSED SYSTEM

The proposed system is built to overcome the drawback of the existing system. The project design involves text to speech. The system receives input after the command the output will get in the form of voice means speech.

4. FLOW CHART

First, when the user starts the system, he uses a microphone to send the input. Basically, what it does is that it takes sound input from the user and it's fed to the pc to process it further. Then, that sound input is fed to the speech to text converter, which converts audio input to text output which is recognizable by the computer and can also be processed by it.

Then that text is parsed and looked for keywords. Our voice command system is built around the system of keywords where it searches the text for keywords to match and when keywords get matched then it gives the relevant output.

This output is in the form of text, which is converted to speech output using a text to speech converter which involves using an optical character recognition (OCR) system. OCR categorizes and identifies the text then the text-to-speech engine converts it to the audio output. This output is transmitted via the speakers which are connected to the audio jack of the raspberry pi as shown in Figure 1.

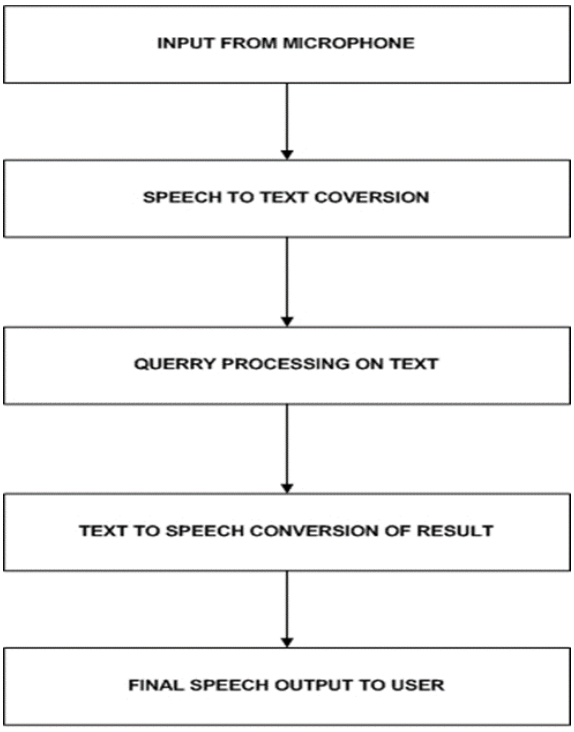

Figure 1. Flow of events in voice command system

5. SYSTEM ARCHITECTURE

First of all the system starts with specific wake words for the voice assistant. Following the wake word, the system proceeds by receiving voice commands from the users with the assistance of a microphone, further analyzing the command and processing the given device will create an appropriate response for the given command. The response is given to the user with the help of speakers as shown in the diagram given below. Users can use the voice assistant for storing and streaming the data. Users can store excess data(Data taking up extra space on the local device) in the wireless storage system of the device with the help of an app on the host's device. This application interacts with the system. Uploading, downloading and streaming of knowledge from the system to the host device is now convenient. The various constraints acting on the system are as follows:- The primary design constraint is the autonomous platform. Since the device is designed to interact with heterogeneous O.S. devices hence, effective GUI and user-friendly interaction are going to be the main design constraint. Creating a user interface that is both interactive and easily navigable is merely important. Also as we are incorporating a personal voice assistant, the need for

effective performance is mandatory. Moreover, Noise cancellation may be a vital constraint into account when voice interaction is employed by the user. Other constraints like processing power also are worth considering. The assistant system is supposed to be quick and responsive even when handling complex instructions, so each feature of the software must be designed and implemented considering the efficiency. As the proposed system involves complex communication algorithms, the wants of all the algorithms must be taken into account for the designated format of input and output generated, their individual working efficiency, and their contribution to overall system application. The system will give the required results as long as the specified requirements are satisfied. The device will need to implement all the specified NLP (Natural language processing) algorithms simultaneously with the server protocols, thus a requirement of synchronized operations is required to hold out the required functionality smoothly for the top users. Moreover, the storage ambiguity needs to be taken care of as multiple users interact with the device simultaneously. The low tolerance of failure of the storage system makes the voice assistant vulnerable in terms of performance. As the number of users increases, a need for higher processing power is required by the device, which could lead to some users getting deprived of their service. Such situations need the implementation of parallel algorithms to meet the needs of the users.

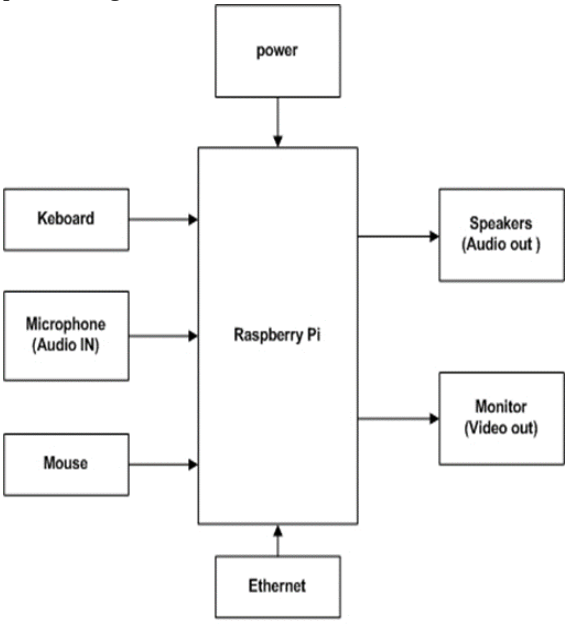

Figure 2. Hardware setup of voice command system

6. COMPONENTS USED

6.1. Raspberry Pi3

Raspberry Pi is the heart of the voice command system as it is involved in every step of processing data to connect components together. The Raspbian operating system is mounted onto the SD card which is then loaded in the card slot to provide a functioning operating system.

Figure 3. Raspberry Pi 3

6.2. MicroSD card 32GB

As RaspberryPi doesn't have its own internal storage. So, a MicroSD card is needed to install RaspberryPi OS.

Figure 4. MicroSD card

6.3. USB wall charger

A USB wall charger is needed to provide power supply to the microcontroller.

Figure 5. USB wall charger

6.4. 3 port USB hub

A 3 port USB hub is needed to convert one USB port into three USB ports.

Figure 6. 3 port USB hub

6.5. USB microphone

The microphone is used to take the audio input of the sound. The audio input passed through the system would be searched for keywords. These keywords are essential for the functioning of the voice command system as modules work by searching for keywords and giving output by matching keywords.

Figure 7. USB Microphone

6.6. Portable speaker with AUX input

Speakers, once the query suggested by the user has been processed, the text output of that question is converted to speech using the web text to speech converter. Now this speech which is that the audio output is distributed to the user using the speakers which are running on audio out.

Figure 8. Speaker with AUX input.

7. WORKING

Figure 9. Poodle: Voice Assistant setup

- Plug into power supply (220V)
- Connect the USB Cable to Raspberry-Pi (5V, 2A)
- Start VNC viewer in the laptop and connect with Raspberry-Pi.

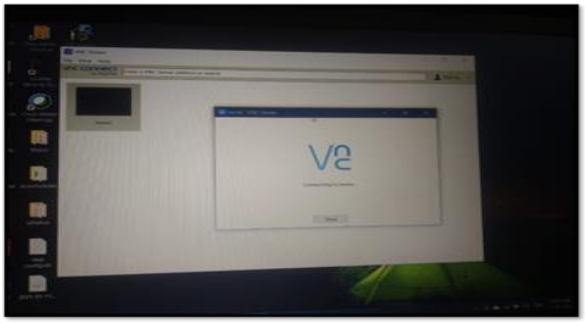

Figure 10. VNC viewer

Open Terminal. After coding the third command, all the 4 bulbs will automatically glow.

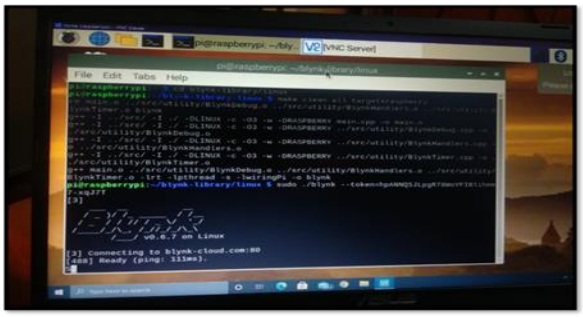

Figure 11. Terminal window of VNC viewer

Figure 12. Automation System

- After setting up the Blynk app.
- Also, we can control the bulbs by giving voice commands through the Google Assistant and also by the Blynk app.

Figure 13. Blynk app

- Users can interact with the Poodle: Voice Assistant using Alexa and manage their day to day task.

8. CONCLUSION

Consumers satisfied with smart speakers were shown to continuously use the service and show cognitive loyalty by remaining attached to the voice assistant provider. The number of people using voice assistants is expected to grow. Voice assistants continuously improve and learn.

Researchers believe that ultimately, the voice assistant might get so smart that it will automatically order a pizza if you say you're hungry. It will use existing data from your previous orders to come to the conclusion that saying you're hungry equals ordering a pizza.

REFERENCES

[1] Veton Këpuska, Gamal Bohouta, "Next-Generation of Virtual Personal Assistants (Microsoft Cortana, Apple Siri, Amazon Alexa and Google Home) ", Published in 2018, ISBN-99978-1-6/18 IEEE.

[2] Sumit kumar Sarda, Yash Shah, Monika Das, Nikita Saibewar, Shivprasad Patil,"VPA: Virtual Personal Assistant" Published in 2017 International Journal of Computer Applications (0975 – 8887) Volume 165 – No 1.

[3] Amit Rakesh Shrivastava, Jayant Gadge," Home Server and NAS using Raspberry Pi", Published in 2017, ISBN 978-6367-3/17 IEEE.

[4] Zhang Huili," Realization of Files Sharing between Linux and Windows based on Samba", Published in 2008 International Seminar on Future BioMedical Information Engineering, ISBN - 978-0-7695-3561- 6/08 IEEE.

[5] Steven Hickson, "Voice Control and Face Recognition on the Raspberry Pi" [online], Available at http://stevenhickson.blogspot.com.

[6] Raspbian[online], Available at http://www.raspbian.org/.

[7] Wayne Wobcke, Anh Nguyen, Van Ho and Alfred Kqzywicki, School of computer science and engineering, University of new south walls, Sydney nsw 2052, Australia, "the smart personal assistant: an overview". (Journal[online])

[8] "Intelligent Personal Assistant", "siri", "cortana", [online] by Wikipedia.

[9] "Pineapple and Raspberry Pi", circuit designs and comparison in google".

BIOGRAPHIES

Dhrati Gupta
UG, Final Year Student
Dept. of ECE,
RKGIT, Ghaziabad, India

Abhishek Dixit
UG, Final Year Student
Dept. of ECE,
RKGIT, Ghaziabad, India

Akshat Mitra
UG, Final Year Student
Dept. of ECE,
RKGIT, Ghaziabad, India

Mr. Anuj Kumar
Assistant Professor
Dept. of ECE,
RKGIT, Ghaziabad, India