Voice Controlled Personal Assistant

Abhinetra Kumar, Aditya Deep, Harshitha P, Chaitanya Anand, Kavya N L

Dept of CS&E, Sapthagiri college of engineering, Bangalore

Abstract- Approx 20% people of the world are suffering from various disabilities, many of them are blind or unable to use their hands effectively. So, in those particular cases the voice recognition system provide a significant help to them, so that they can give certain commands to computer and get things done through voice input. This project provides supports for these disabled humans. Python AI project is a Voice activated Raspberry pi assistant. It is not exactly an AI but it can play music, get weather forecast, do some simple conversion, send emails, solve mathematic calculations, Wikipedia, take attendance based on voice, control projector using voice command.

Keywords- Python, Raspberry Pi, Speech recognition.

I. INTRODUCTION

On a high rate, we have many new technologies available to everyone at our disposal. For that, low cost programmable hardware is largely available to people and used widely. The important aspect is that it can be easily programmed relatively from an environment already developed like Raspbian OS like **Raspberry Pi**.

And now-a-days internet facilities are available at cheap rate and there are a lot of tools and APIs that enable many of the functionalities and as well provide personal information on different platforms such as Google personal information and other sources. This project arises from the idea of intelligent personal assistants like Siri, Google Now, and Cortana which are all smartphone and PC features. And this idea moved on to the rise of intelligent personal assistants for a vehicle or a car. It is being developed based on open source code having in mind the above considerations, which could be modified or extended or further developed by a larger community of software developers.

Raspberry Pi is known as the credit-card sized computer which was initially designed for education, inspired by the 1981 BBC Micro. It is a low-cost device through which various applications of IoT can be developed. Knowledge of an Operating System is a must to work with a Raspberry Pi board. With inbuilt Wi-Fi and Bluetooth connectivity, it makes Raspberry Pi the most suitable hardware to work on.

It runs on a Quad core ARM Cortex A-53 at a clock speed of 1.2GHz. It has a 1 Gigabyte DDR2 RAM with a clock speed of 900MHz.

Fig.1 - Raspberry Pi 3 Model

II. RELATED WORKS IN THE FIELD

With an idea of implementing the voice to achieve solution to problems has open several possible doors for the technology to work on it and make it as one of the successful ideas for physically disabled human.

Many well-known companies such as including Google, Amazon, Apple, Microsoft have worked on the idea of voice assistant in the form of Google Now, Alexa, Siri and Cortana respectively. The primary aim of these system is to make the human to computer interaction easier by making it more like human to human interaction by processing the voice input. All these systems are available with lots of features making it really an expensive thing for an average man thus, making it unaffordable by everyone.

Several years back, Mark Zuckerberg tried to develop an Artificially Intelligent personal assistant named 'Jarvis' inspired from the movie 'Iron Man'. He quite succeeded in the project taking too much efforts to code in order to make it possible. These are some of the related works in the field which inspired us to start with this project.

III. PROPOSED SYSTEM

This work focuses on taking voice as an input and after processing the speech according to command it provides the output in the form of speech. It is designed in such a way that the user using it feels like he or she is actually talking to a real person instead of a system.

A. Architecture

Hardware Implementation

Raspberry Pi is a single board computer that works on LINUX operating system. Printed circuit board has tons of features and many ports so that it can run multiple applications and at the same time multiple hardware can be connected on which those applications run. The Raspbian OS is installed on SD card which will be loaded in the card slot to give the functional OS for running some instructions.

Speakers are used here to give the output of the processed voice which is given by the user. The output will be the information regarding the query, the text as an output of this query is converted to speech using the online text to speech converter module.

Microphone is used as the input provider to this system wherein the user will provide his or her command through this microphone which later can be fed as an input to system for further processing.

Wi-Fi is wireless fidelity which is used to connect to the internet wirelessly. It also provides high speed data transferring and this is used here because our project requires continuous internet connectivity to perform all operations.

Projector is used to project the ppts and previous and forward option can be managed by voice command.

Power given to the main hardware of system is used to activate the functioning of it.

Ethernet is used in this proposed system to collect various information according to the request provided by the user.
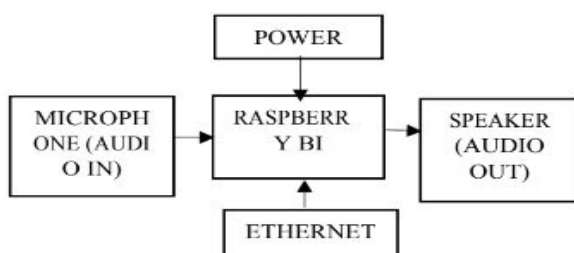

Fig.2 - Architecture

Modules Implemented

The main modules implemented int his systems are given below

- Speech-To-Text (STT) engine
- Text-To-Speech (TTS) engine

The additional module used in this system are for various purpose listed below:

- Wolfram Alpha
- Wikipedia
- Pyttsx3
- Speech Recognition
- Datetime
- Webbrowser
- Smtplib

Speech to Text Engine

(AVS) is a Speech-To-Text (STT) engine, is used to convert the commands given by the user in audio input to text form, so that these commands can be interpreted by the modules properly. To use (AVS) engine, an application has to be created in the Amazon developers console and the generated API key has to be used to access the speech engine. It requires continuous internet connection as data is sent over the Amazon servers.

Text to Speech Engine

(AVS) is Text-To-Speech (TTS) engine which is used to create a sound version of the text in a computer. TTS enables the speaking or reading of computer display for the vision lack peoples, or may simply be used to augment the reading of a text message. To use (AVS) engine, an application has to be created in the Amazon developers console and the generated API key has to be used to access the speech engine. It requires continuous internet connection.

Wolfram Alpha

Wolfram-Alpha is python module. It is a search engine for computing answers and providing knowledge. It mainly searches some categories like mathematics, science and technology, society and culture and everyday life. Computation is done from these categories only.

Pyttsx3 : pyttsx3 is a text-to-speech conversion module in python. Other modules which are responsible for converting text into speech sometimes requires net connection but pyttsx3 can work offline also, and is compatible with both versions of python i.e., 2 and 3.

Speech Recognition

Speech-recognition is a python module that is used for recognising the speech which is spoken by the user so that it will be converted in appropriate text and which will be useful for searching net easily.

Datetime

In python, datetime is the module which deals with date, times and even time interval also. It is very useful for the user to send mails or doing something on accurate time interval.

Wikipedia

Wikipedia is a python module or library which is responsible for accessing the data or information directly from Wikipedia. Wikipedia wraps the Media-Wiki API.

Webbrowser

The Web browser is python module which is used to deal with the browsers to show the contents or information searched from the internet in the HTML format on different web browsers.

Smtplib

Smtplib is a python module, which defines an SMTP client session object so that the user can send mails to another receiver via internet. It is very useful for attachments which can be send easily along with normal text information.

IV. WORKING OF PROPOSED SYSTEM

```
┌─────────────────────────────────┐
│     INPUT FROM MICROPHONE       │
└─────────────────────────────────┘
                 │
                 ▼
┌─────────────────────────────────┐
│   SPEECH TO TEXT CONVERSION     │
└─────────────────────────────────┘
                 │
                 ▼
┌─────────────────────────────────┐
│   QUERY PROCESSING ON TEXT      │
└─────────────────────────────────┘
                 │
                 ▼
┌─────────────────────────────────┐
│  TEXT TO SPEECH CONVERSION OF   │
│            RESULT               │
└─────────────────────────────────┘
                 │
                 ▼
┌─────────────────────────────────┐
│   FINAL SPEECH OUTPUT TO USER   │
└─────────────────────────────────┘
```

Fig.3 – Working of the System

Input from Microphone

Users provides its request through a microphone by stating its demand or requests in the form of speech. The important point which one must take care in this part is that the microphone used in the system must possess properties which makes it able to capture the sound from user effectively. The microphone failing to satisfy this property may results in less efficient system.

Speech to text conversion

Once the user provides the input through microphone, the speech needs to be converted to text in order for the system to understand. This conversion of speech to text is achieved by using speech recognition module in python.

Query Processing on Text

Once the conversion of speech to text is performed, the text containing the request of user is processed to provide the outcome of request. Here, different modules used in this system comes into play. The result may be retrieved using any of the possible modules from the one listed above in section [III].

Text to Speech Conversion of Result

Since the results after processing the query are in the form of text and the required outcome needs to be in the form of speech, the results required to be converted into speech. Thus, it makes use of pyttsx3 module from python.

Final Speech Output

After the conversion of result into speech, it is delivered to user from the speaker hardware connected to system.
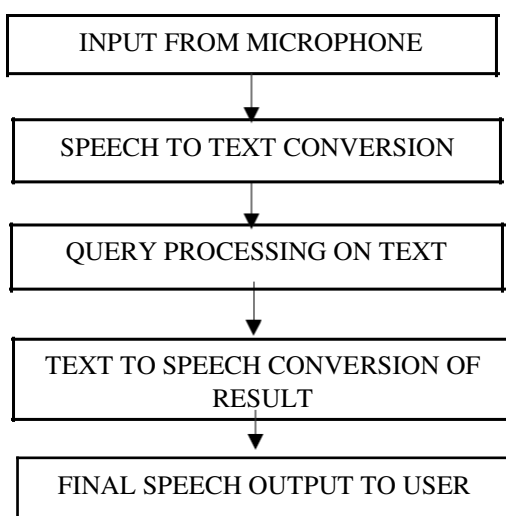
V. RESULT

Voice Controlled Personal Assistant is designed to work on the speech request of the users through the microphone. The words contained in the speech command is searched from storage in the python code and respective actions are taken. System response to command given by users like searching a particular thing from internet or Wikipedia. It also performs solving of equation mathematically to find the value of a variable. It can perform simple arithmetic operations. In addition ,it sends emails along with different type of attachment such as pdf, images etc. The system also can predict the weather forecast of a particular place, can play music and greet you according to the time of use of the system.

VI.   CONCLUSIONS

The Voice Controlled Personal Assistant system is developed in python programming language by making use of most of module present in the language. The system can be used efficiently by almost everyone who can not afford to use the system developed by Google, Amazon and so on due to its high price. Moreover, one can also make changes into it as per their requirement thus making it an open source product.

VII.   FUTURE WORK

The system requires the internet connectivity to be proper for normal functionality of this system. But considering the speed of internet it is not necessary that everyone receives the same speed connectivity thereby can affect the system performance. So, one solution to this problem can be the development of similar product which can work offline. Another work which can be included in this proposed system can be addition of native languages which may result in the more demand of the current system.

VIII.   REFERENCES

[1]   Manoj Kumar Sharma, "Speech Recognition: A Review", School of Engineering & Technology, Aug 2015.

[2]   Kishore Kumar R, "A python based virtual assistant using raspberry pie", Saveetha engineering college,Vol 5, July 2018.

[3]   Emad S, "Voice controlled personal assistant using raspberry pie",Vol 8, Nov 2017.

[4]   Saira Shamsudheen K.S, Liya Mol Aliyar, "Intelligent Personal Assistant for Vehicles", Jul 2018.

[5]   Saira Shamsudheen, "Intelligent personal assistant for vechiles", Illahi College of Engineering, Vol 5, Jul 2017.