Patuakhali Science and Technology University

Project Report On

The Voice Enabled Personal Assistant for PC

Course Code: CCE-310

Course Title: Software Development Project-1

Submitted By

Name: Mehedi Hasan Rabbi

ID: 1802052

Reg. No: 08462

Level: 03

Semester: 01

Submitted To

Professor Dr. Md Abdul Masud

Department of Computer Science and Information Technology

Patuakhali Science and Technology University

Assistant Professor Arpita Howlader

Department of Computer and Communication Engineering

Patuakhali Science and Technology University

Submission Date: 22 January,2023

DECLARATION

I declare that this report presented in it is my own and has been generated by me as the result of my work.

 I confirm that:

- This Work is done wholly or mainly while in candidature for a degree at this University.

- This report has not been previously submitted for any degree at this university or any other educational institutes.

- I have quoted from the work of others; the source is always given. With the exception of such quotations, this report is entire our own work.

………………………………………

Mehedi Hasan Rabbi

ABSTRACT

Personal Assistants, or conversational interfaces, or chat bots reinvent a new way for individuals to interact with computes. A Personal Virtual Assistant allows a user to simply ask questions in the same manner that they would address a human, and are even capable of doing some basic tasks like opening apps, reading out news, taking notes etc., with just a voice command. Personal Assistants like Google Assistant, Alexa, Siri works by Speech Recognition (Speech-to-text) and Text-to-Speech.

Keywords: Personal Assistants; JARVIS; conversational interfaces; Speech Recognition; Text-to-Speech.

ACKNOWLEDGEMENTS

Table of Contents

1. INTRODUCTION

In today's era almost, all tasks are digitalized. We have Smartphone in hands and it is nothing less than having world at your fingertips. These days we aren't even using fingers. We just speak of the task and it is done. There exist systems where we can say Text Dad, "I'll be late today." And the text is sent. That is the task of a Virtual Assistant. It also supports specialized task such as booking a flight, or finding cheapest book online from various ecommerce sites and then providing an interface to book an order are helping automate search, discovery and online order operations. Virtual Assistants are software programs that help you ease your day to day tasks, such as showing weather report, creating reminders, making shopping lists etc. They can take commands via text (online chat bots) or by voice. Voice based intelligent assistants need an invoking word or wake word to activate the listener, followed by the command. For my project the wake word is JARVIS. We have so many virtual assistants, such as Apple's Siri, Amazon's Alexa and Microsoft's Cortana. For this project, wake word was chosen JARVIS.

This system is designed to be used efficiently on desktops. Personal assistant software improves user productivity by managing routine tasks of the user and by providing information from online sources to the user. JARVIS is effortless to use. Call the wake word 'JARVIS' followed by the command. And within seconds, it gets executed. Voice searches have dominated over text search. Web searches conducted via mobile devices have only just overtaken those carried out using a computer and the analysts are already predicting that 50% of searches will be via voice by 2022.Virtual assistants are turning out to be smarter than ever. Allow your intelligent assistant to make email work for you. Detect intent, pick out important information, automate processes, and deliver personalized responses. This project was started on the premise that there is sufficient amount of openly available data and information on the web that can be utilized to build a virtual assistant that has access to making intelligent decisions for routine user activities.

2. PROJECT OVERVIEW

In this proposed concept effective way of implementing a Personal voice assistant, Speech Recognition library has many in-built functions, that will let the assistant understand the command given by user and the response will be sent back to user in voice, with Text to Speech functions. When assistant captures the voice command given by user, the under lying algorithms will convert the voice into text. And according to the keywords present in the text (command given by user), respective action will be performed by the assistant. This is made possible with the functions present in different libraries. Also, the assistant was able to achieve all the functionalities with help of some API's.

3. OBJECTIVE OF THE PROJECT

Main objective of building personal assistant software (a virtual assistant) is using semantic data sources available on the web, user generated content and providing knowledge from knowledge databases. The main purpose of an intelligent virtual assistant is to answer questions that users may have. This may be done in a business environment, for example, on the business website, with a chat interface. Virtual assistants can tremendously save you time. We spend hours in online research and then making the report in our terms of understanding. One of the main advantages of voice searches is their rapidity. In fact, voice is reputed to be four times faster than a written search: whereas we can write about 40 words per minute, we are capable of speaking around 150 during the same period of time. In this respect, the ability of personal assistants to accurately recognize spoken words are a prerequisite for them to be adopted by consumers.

Other objectives of the project included:

- Practicality

- Efficiency

- Cost

- Flexibility

- Portability

4. REVIEW ON RELATED WORK

AbhayDekate (2016) et al. presented in the Modern Era of fast moving technology we can do things which we never thought we could do before but, to achieve and accomplish these thoughts there is a need for a platform which can automate all our tasks with ease and comfort. The Hardware device captures the audio request through microphone and processes the request so that the device can respond to the individual using in-built speaker module. For Example, if you ask the device 'what's the weather?' or 'how's traffic?' using its built-in skills, it looks up the weather and traffic status respectively and then returns the response to the customer through connected speaker.

Dr.Kshama V. Kulhalli (2017) et al. proposed the Most famous application of iPhone is "SIRI" which helps the end user to communicate end user mobile with voice and it also responds to the voice commands of the user. Same kind of application is also developed by the Google that is "Google Voice Search" which is used for in Android Phones. But this Application mostly works with Internet Connections. But our Proposed System has capability to work with and without Internet Connectivity.

VetonKëpuska (2018) proposed one of the goals of Artificial intelligence (AI) is the realization of natural dialogue between humans and machines. in recent years, the dialogue systems, also known as interactive conversational systems are the fastest growing area in AI. Many companies have used the dialogue systems technology to establish various kinds of Virtual Personal Assistants(VPAs) based on their applications and areas, such as Microsoft's Cortana, Apple's Siri, Amazon Alexa, Google Assistant, and Facebook's M. However, in this proposal, we have used the multi-modal dialogue systems.

Isha S. Dubey (2019) et al. proposed about a different combination of a reading machine (OCR), virtual assistant and Domotics system using Raspberry-Pi which will be a combination of a great system. This is a helpful aid for visually impaired people and people with disabilities. OCR stands for optical character recognition where it recognizes the present text and converts them into audio speech using pre and post processing with gTTS (Google Text to Speech).

TusharGharge (2019) et al. presented the problem of user while developing a computer program. Developing a computer program is not an easy task it needs hardware resources which user have to handle. While continuous typing the code there may be possibility of injuries to the fingers of the user.

5. FEATURES, TECHNOLOGY, TOOLS, LIMITATION

5.1 Features
- Tell time and date
- Switch between male and female voice
- Voice command and text command mode
- Send WhatsApp message using voice
- Wikipedia search
- Google and YouTube search
- Give weather information
- Open file explorer, chrome and other application
- Read any selected text
- Tell a joke
- Take screenshot
- Take note
- Password generate
- Flip a coin or roll a die
- Give system usage information and more

5.2 Technology Used
- Python
- 'pyttsx3' – Python text-to-speech library
- 'speech_recognition' – a google speech recognition API.
- Natural Language Toolkit - NLTK is a leading platform for building Python programs to work with human language data.

- 'PyAutoGUI' - cross-platform GUI automation Python module for human beings.
- 'weather API' – Simple and fast and free weather API from OpenWeatherMap

5.3 Tools Used

As the functional requirements, we have indicated:

1) Development tools, Visual Studio Code and a sustainable environment for development.
2) Python interpreter
3) Different types of python module and library.

5.4 Limitation

1) Security is somewhere an issue, there is no voice command encryption in this project.
2) Background voice can interfere.
3) Misinterpretation because of accents and may cause inaccurate results.

6. WORKING PROCEDURE/METHODOLOGY

The assistant, on starting, will initially wait for the input to be given from user. If the user gives input command, via voice, the assistant will capture it, and searches for the keyword present in the input command. If the assistant was able to find a key word, then it will perform the task accordingly, and returns the output back to user, in voice. If not, the assistant will again start waiting for the user to give input. Each of these functionalities are having their own importance in the whole system working.

The working of a Virtual Assistant uses the following principles:

1. **Natural Language Processing:** Natural Language Processing (NLP) refers to an AI approach to intelligent programming using a natural language such as English. In our project we have use it as tokenizer.
2. **Automatic Speech Recognition:** To understand commands according to the user's input.
3. **Text Analyzing:** Converted text is just letters for computers. A piece of library converts text to something understandable for the computer. Computer understands the command, so Virtual Assistants converts this text to a computer command.
4. **Check Command:** If the assistant was able to find a key word, then it will perform the task accordingly, and returns the output back to user, in voice. If not, the assistant will again start waiting for the user to give input. Each of these functionalities are having their own importance in the whole system working.
5. **Execute:** If the given task is in range of the virtual assistant then it will done by the virtual assistant and wait for the next command.
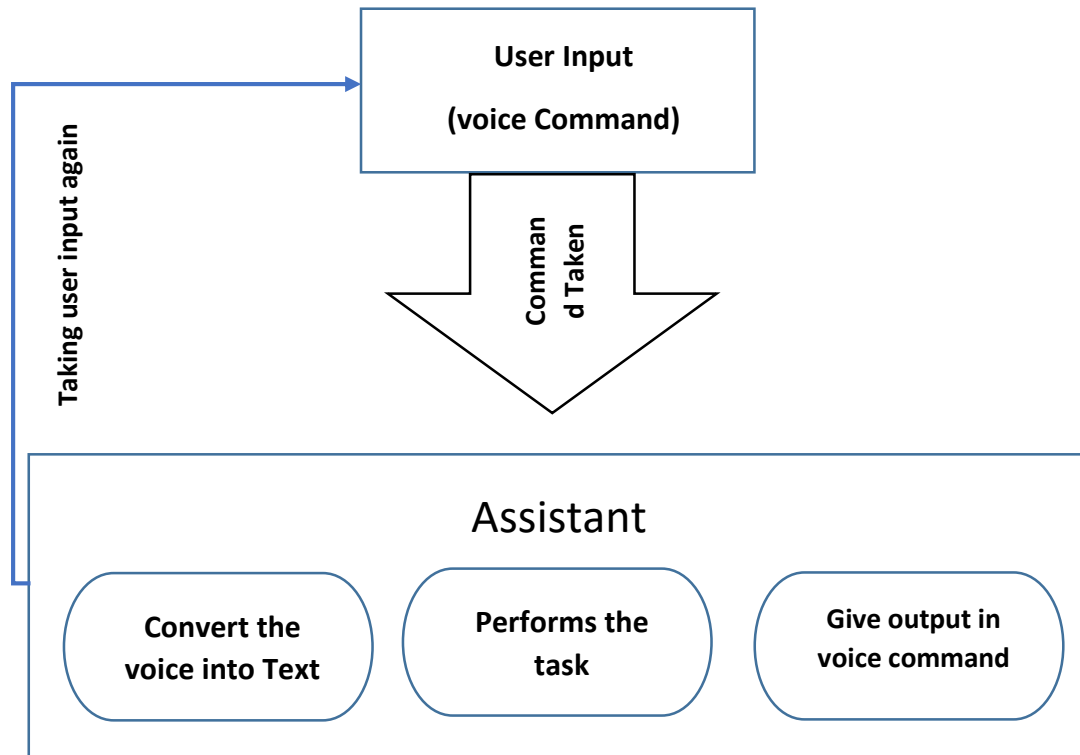
6.1 Block Diagram

User Input

(voice Command)

Taking user input again

Command Taken

Assistant

Convert the voice into Text

Performs the task

Give output in voice command

Fig: Block Diagram

6.2 Data Flow Diagram

A. DFD Level 0 (Context Level Diagram):

Send Request

User

Virtual Assistant

Receives Response
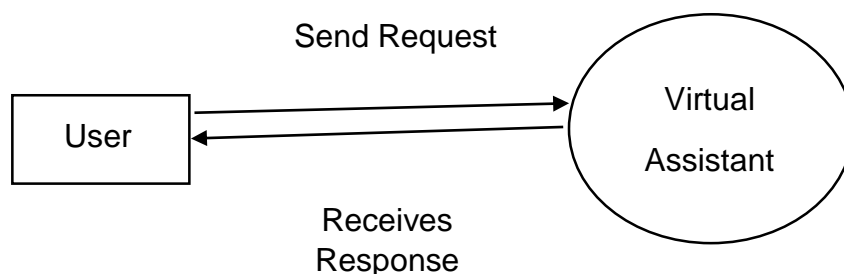
Fig: Data Flow Diagram L0

B. DFD Level 1:

1.0
Assistance

↓

2.0
User Data

↓

3.0
Setting
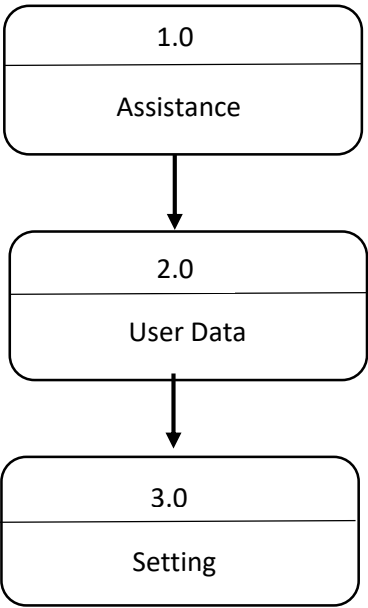
Fig: Data Flow Diagram L0

6.3 Use Case Diagram

Virtual Assistant

+queries
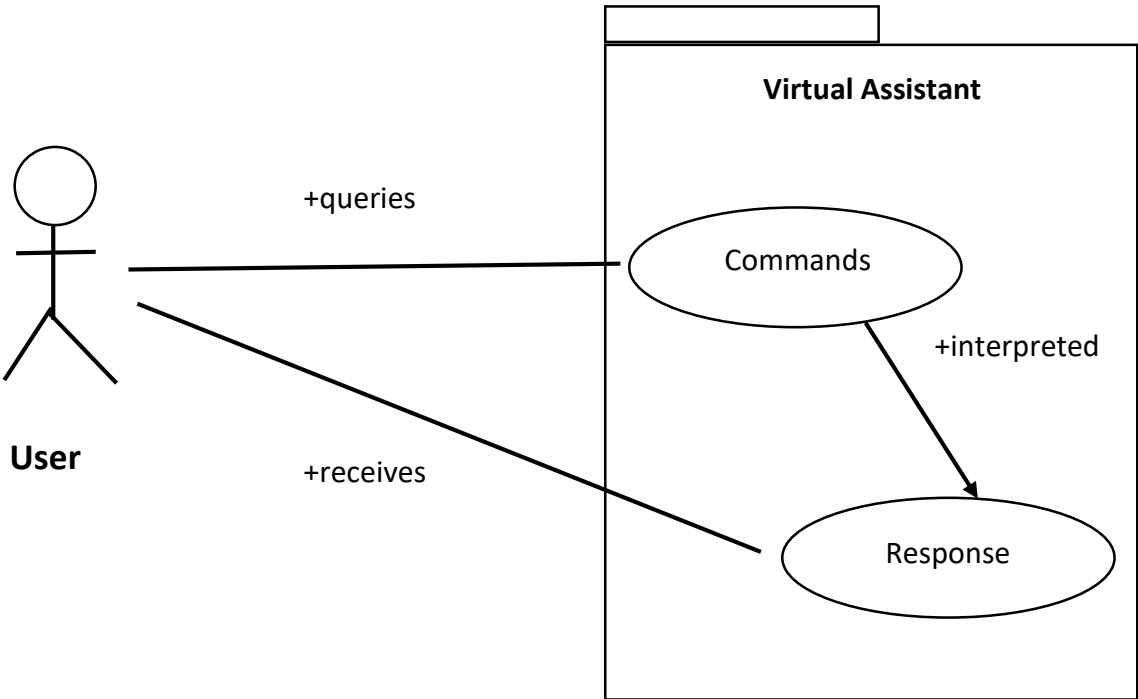
Commands

+interpreted

User

+receives

Response

Fig: Use Case Diagram

7. EXPERIMENTAL SETTING INCLUDING TOOLS & LANGUAGE

To manage and do experiment on this project one have install all the requirement which are given below:

Hardware Requirements: The following are the minimum necessary hardware requirements necessary for the proper implementation of the Virtual Assistent System:

- A 32-bit 2.2GHz processor

- Windows 7 and upwards operating system or Linux or macOS

- 2 GB RAM processor

Language: Install python 3.9.x or upper version.

Modules and APIs:

1. 'pyttsx3' – Python text-to-speech library
2. 'speech_recognition' – a google speech recognition API.
3. Natural Language Toolkit - NLTK is a leading platform for building Python programs to work with human language data.
4. 'PyAutoGUI' - cross-platform GUI automation Python module for human beings.
5. 'weather API' – Simple and fast and free weather API from OpenWeatherMap

8. WORKING MODEL

• User Input—The assistant will wait for the user to give voice command for further processing.

```
PS C:\Users\Mehedi Hasan\Desktop\J.A.R.V.I.S> python -u "c:\Users\Mehedi Hasan\Desktop\J.A.R.V.I.S\Jarvis.py"
Listening...
Recognizning...
how are you Jarvis
Listening...
Recognizning...
```

Fig: User Input

• Introducing to user—The user who is asking assistant to introduce itself, will display the following.

```
PS C:\Users\Mehedi Hasan\Desktop\J.A.R.V.I.S> python -u "c:\Users\Mehedi Hasan\Desktop\J.A.R.V.I.S\Jarvis.py"
Allow me to introduce myself I am Jarvis, the virtual artificial intelligence and I'm here to assist you with a variety of tasks as best I
can, 24 hours a day seven days a week.
Listening...
```

Fig: Introducing itself

• Taking a sample note—If the user has a small note to be taken, he can ask the assistant to do so, and the assistant will take the notes and save it in a notepad file.

Fig: Taking short note

• Showing Note—If the user asks the assistant to display the note, and to speak out the note, the assistant will do so.

Fig: Showing Note

• Sending WhatsApp message—If the user asks the assistant to send a message through WhatsApp, the assistant will do so.
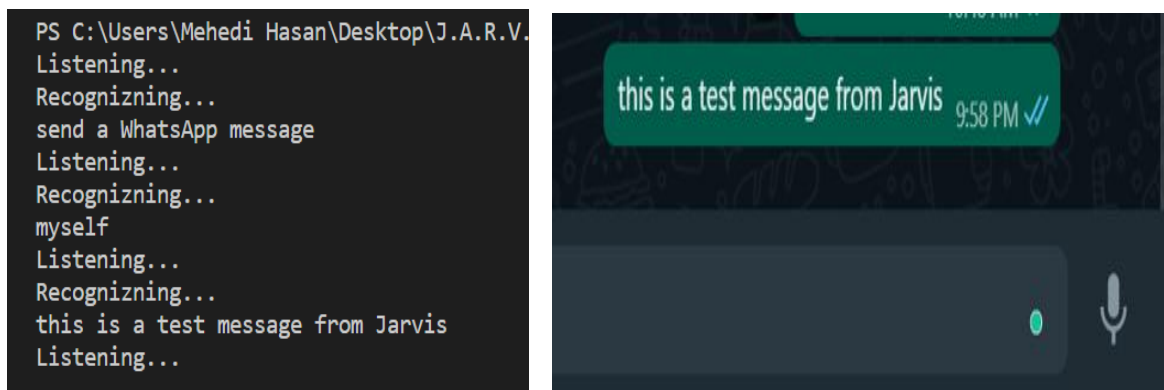

Fig: Sending WhatsApp message

• YouTube searches—If the user asks the assistant to do some YouTube searches, the assistant will do that. It will ask the user, what to search in YouTube. After receiving the input, it will open the YouTube page with that respective search.
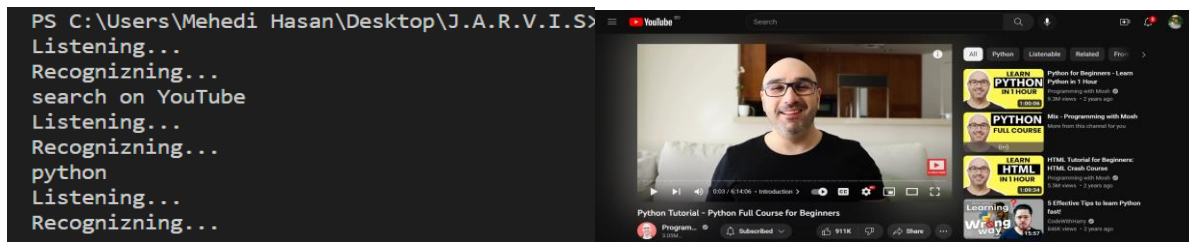
Fig: Asking assistant to open YouTube

•Taking Screenshots—If the user asks the assistant to take screenshots, the assistant will take a screenshot, and will save it in a particular location.
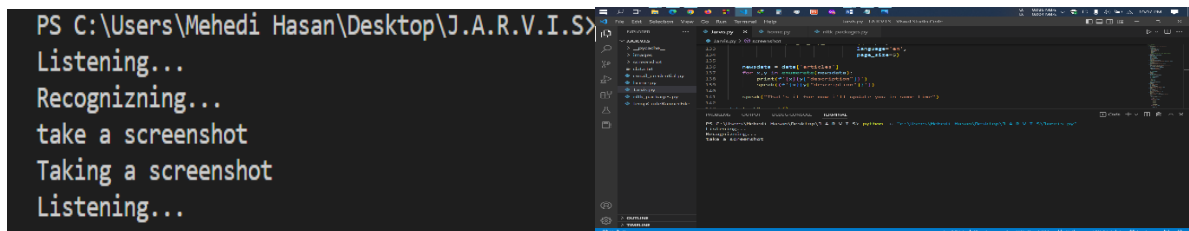

Fig: Assistant taking a screenshot

•Weather Information—If the user asks the assistant to give weather information, the assistant will give information of given location.
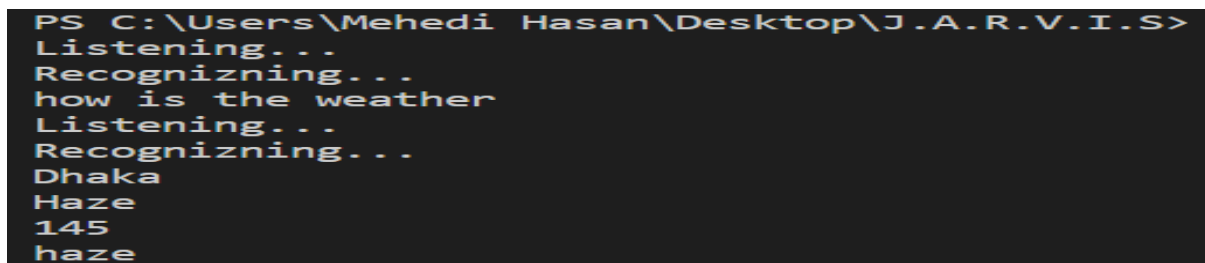

Fig: Weather information of Dhaka

• Opening Applications—If the user asks the assistant to open an application, like MS Word, or any other, the assistant will do so immediately. And also, it will speak that it opens the application.
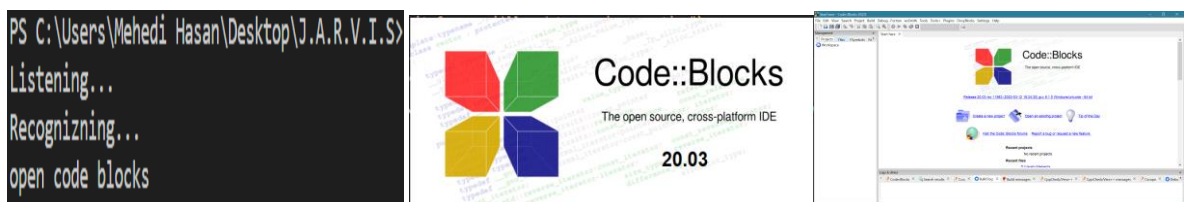

Fig 9: Assistant opening Code::Blocks

Report on Voice Enable Virtual Assistant

9. Application of the project

Virtual assistants typically perform simple jobs for end users, including the following:

- ❖ Home automation
- ❖ System automation
- ❖ This type of assistant program is helpful for blind persons.
- ❖ Administrative work
- ❖ Social media production
- ❖ Content management
- ❖ Research partner
- ❖ Adding tasks to a calendar;
- ❖ Providing information that would normally be searched in a web browser;
- ❖ Controlling and checking the status of smart home devices, including lights, cameras and thermostats;
- ❖ Making and receiving phone calls;
- ❖ Creating text messages;
- ❖ Getting directions;
- ❖ Hearing news and weather reports;
- ❖ Finding hotels or restaurants;
- ❖ Checking flight reservations;
- ❖ Listening to music; and
- ❖ Playing games.

10. 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.

11. FUTURE WORK

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.

12. REFERENCE

[1] www.stackoverflow.com

[2] www.tutorialspoint.com

[3] www.google.com

[4] www.python.org

[5] https://pypi.org/

[6] https://cloud.google.com/speech-to-text/

[7] https://geeksforgeeks.org/

[8] https://openweathermap.org/api

[9] https://www.nltk.org/