**Project Id-**

**A**

**Synopsis**

**on**

**VOICE RECOGNITION MODEL USING PYTHON**

**Submitted in partial fulfilment of the requirements** **for the award**

**of the degree of**

**Bachelor of Technology**

**in**

**Computer Science and Engineering**

**By**

**GOVIND NAIR (1680210456)**

**SATENDRA PAL (1680210031)**

**VICKY MADDHESHIYA (1680210040)**

**MOHIT KUMAR GUPTA (1680210022)**

**Under the Supervision of**

**Mr NK SHARMA SIR**

**NITRA Technical Campus**

**Ghaziabad 201002**

**Affiliated to**



**Dr. A.P.J. Abdul Kalam Technical University**

**Lucknow**

**INDEX**

|  |  |  |
| --- | --- | --- |
| **S.No.** | **HEADING** | **Page No.** |
| **1.** | **INTRODUCTION** | **3** |
| **2.** | **BRIEF LITERATURE SURVEY** | **4-5** |
| **3.** | **PROBLEM FORMULATION** | **9** |
| **4.** | **OBJECTIVES** | **10** |
| **5.** | **METHODOLOGY** | **11** |
| **6.** | **REFERENCE** | **12** |

**INTRODUCTION**

The basic idea behind this project is to create a simple stand-alone application that helps less tech savvy people in the world to use the computer without feeling ignorant or computer illiterate. Computers have become a very important devices and as well as less expensive over time. The application works same like Cortana/Siri/ Google Assistant etc. But the application deals with the computer itself mainly. The U.I of the application is self-explanatory and minimal. Currently it takes text as input as most of the people are not very good at speaking.



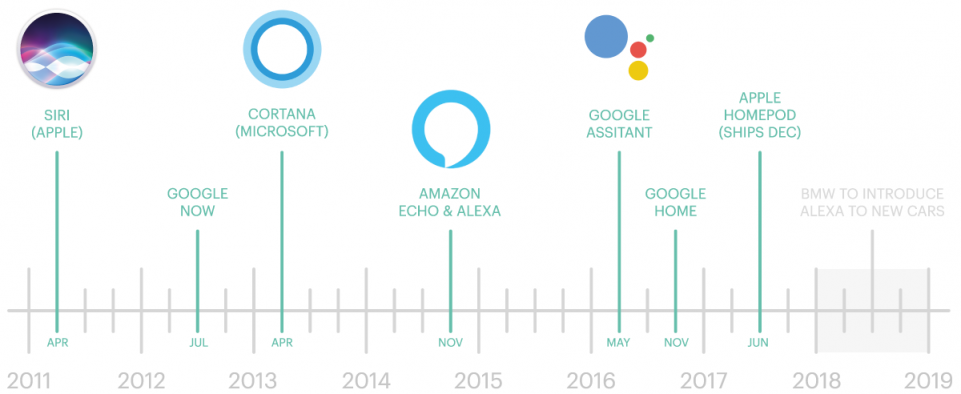
***What is a Voice Assistant* ?**

A voice assistant or intelligent personal assistant is a software agent that can perform tasks or services for an individual based on verbal commands i.e. by interpreting human speech and respond via synthesized voices. Users can ask their assistants’ questions, control home automation devices, and media playback via voice, and manage other basic tasks such as email, to-do lists, open or close any application etc. with verbal commands.

Let me give you the example of braina AI which is an intelligent personal assistant, human language interface, and automation and voice recognition software for Windows PC. Braina is a multi-functional AI software that allows you to interact with your computer using voice commands in most of the languages of the world. Braina also allows you to accurately convert speech to text in over 100 different languages of the world.

**BRIEF LITERATURE SURVEY:**

***History of voice assistant* :**

****

In recent times, Voice assistants got the major platform after Apple integrated the most astonishing Virtual Assistant — Siri which is officially a part of Apple Inc. But the timeline of greatest evolution began with the year 1962 event at the Seattle World Fair where IBM displayed a unique apparatus called Shoebox. It was the actual size of a shoebox and could perform scientific functions and can perceive 16 words and also speak them in the human recognizable voice with 0 to 9 numerical digits.

During the period of the 1970s, researchers at Carnegie Mellon University in Pittsburgh, Pennsylvania — with the considerable help of the U.S Department of Defense and Defense Advanced Research Projects Agency (DARPA) — made Harpy. It could understand almost 1,000 words, which is approximately the vocabulary of a three-year-old child. Big organizations like Apple and IBM sooner in the 90s started to make things that utilized voice acknowledgment. In 1993, Macintosh began to building speech recognition with its Macintosh PCs with Plain Talk. In April 1997, Dragon NaturallySpeaking was the first constant dictation product which could comprehend around 100 words and transform it into readable content.

The basic idea behind this project is to create a simple stand-alone application that helps less tech savvy people in the world to use the computer without feeling ignorant or computer illiterate. Computers have became a very important device and as well as less expensive over time.

The application works same like Siri/ Google Assistant etc. But the application deals with the computer itself mainly. The U.I of the application is self-explanatory and minimal. Currently it takes text as input as most of the people are not very good at speaking.

***Conceptual framework*:**

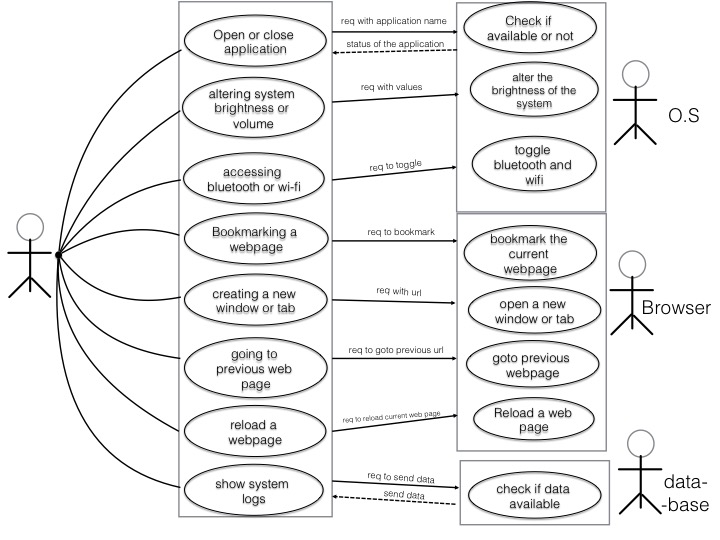
There is always scope for improvement. So the application was built using M.V.C architecture, incremental process design. These design structures follow the modular approach which makes it easy to add in new features to the system. The system also implements the singleton pattern and the single responsibility principle which ensure the individual functioning of the modules.

***Functions*:**   
The system provides all the below stated functions.

* displaying system information such as network usage, ram etc.
* opening and closing an application
* altering system brightness, volume
* accessing Bluetooth, WIFI
* bookmarking a webpage
* creating a new window or a tab in browser
* going to the previous webpage on the web browser
* reloading a webpage on the web browser
* Keylogging with keystrokes segregated according to the application and time stamps.
* browser history tracking

**UML DIAGRAMS:**

***Use Case Diagram*:**   
A use case is a set of scenarios that describe an interaction between a user and a system.  A use case diagram displays the relationship among actors and use cases.  The two main components of a use case diagram are use cases and actors.

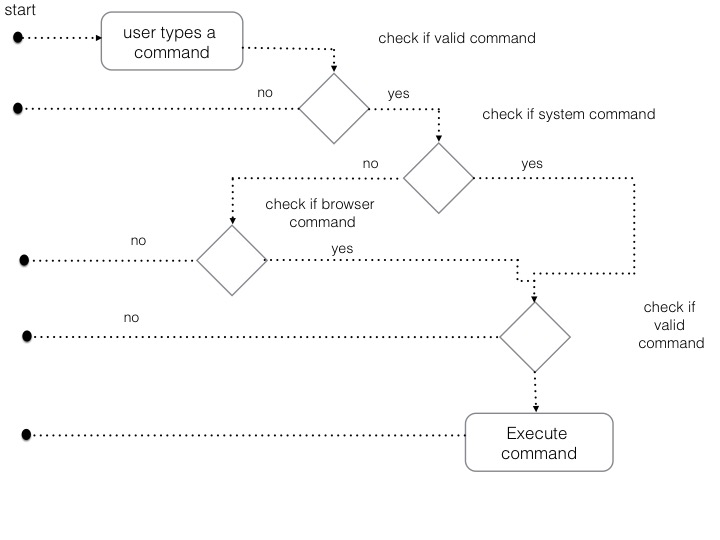
****

***Activity Diagram*:**

Activity diagram is basically a flowchart to represent the flow from one activity to another activity. The activity can be described as an operation of the system. The control flow is drawn from one operation to another. This flow can be sequential, branched, or concurrent. Activity diagrams deal with all type of flow control by using different elements such as fork, join, etc.

**USER INPUT IS CONVERTED INTO TEXT BY USING GOOGLE SPEECH TO TEXT.**

**AND PROCESSED FURTHER:**

****

***Tools Used*:**

The project is mainly built using the Python programming language according to the required functionalities to attain lightning speed and smooth execution. Various python libraries must be implemented to attain various functionalities of text to speech and speech to text.

### ***The Future of Voice Assistants :***

****

To call any technology that makes our lives easier by one name is almost impossible. There are a variety of terms that refer to agents that can perform tasks or services for an individual, and they are almost interchangeable but not quite. They differ mainly based on how we interact with the technology, the app, or a combination of both. Here are some basic definitions, similarities, and differences:

* ***Intelligent Personal Assistant***: This is software that can assist people with basic tasks, usually using natural language. Intelligent personal assistants can go online and search for an answer to a user’s question. Either text or voice can trigger an action.
* ***Automated Personal Assistant***: This term is synonymous with intelligent personal assistant.
* ***Smart Assistant***: This term usually refers to the types of physical items that can provide various services by using smart speakers that listen for a wake word to become active and perform certain tasks. Amazon’s Echo, Google’s Home, and Apple’s Home Pod are types of smart assistants.
* ***Virtual Digital Assistants***: These are automated software applications or platforms that assist the user by understanding natural language in either written or spoken form.
* ***Voice Assistant***: The key here is voice. A voice assistant is a digital assistant that uses voice recognition, speech synthesis, and natural language processing (NLP) to provide a service through a particular application.

The number of people using voice assistants is expected to grow. According to the Voicebot Smart Speaker Consumer Adoption Report 2019, almost 5 percent of people who do not own a smart speaker plan to purchase one. If this holds true, the user base of smart speaker users will grow 70 percent, meaning a quarter of adults in the United States will own a smart speaker.

**NEED AND SIGNIFICANCE**

* ***Provide information such as weather, facts from e.g.***[***Wikipedia***](https://en.wikipedia.org/wiki/Wikipedia)***or***[***IMDb***](https://en.wikipedia.org/wiki/IMDb)***, set an alarm, make to-do lists and shopping lists***
* ***Play music from streaming services such as***[***Spotify***](https://en.wikipedia.org/wiki/Spotify)***and***[***Pandora***](https://en.wikipedia.org/wiki/Pandora_Radio)***; play radio stations; read***[***audiobooks***](https://en.wikipedia.org/wiki/Audiobooks)
* ***Play videos, TV shows or movies on televisions, streaming from e.g.***[***Netflix***](https://en.wikipedia.org/wiki/Netflix)
* [***Conversational commerce***](https://en.wikipedia.org/wiki/Conversational_commerce)***(see below)***
* ***Assist public interactions with government (see***[***Artificial intelligence in government***](https://en.wikipedia.org/wiki/Artificial_intelligence_in_government)***)***
* ***Complement and/or replace customer service by humans. One report estimated that an automated online assistant produced a 30% decrease in the work-load for a human-provided***[***call centre***](https://en.wikipedia.org/wiki/Call_centre)***.***

**OBJECTIVES**

* **Developing an efficient PC voice assistant.**
* **Understanding the concept of Natural Language Process through this project.**
* **This project helps in understanding the concepts of data science.**
* **Prepare and optimize for voice success.**
* **Foster trust through consistency.**
* **Priorities and developing a voice assistant with the least possible complexity.**
* **Configuring the Language input & output as English.**
* **To understand the working of Voice assistant through various Natural Language Process.**
* **Decreasing the complexity of understanding the Voice Input from the user.**

**METHODOLOGY**

* ***The Voice assistant Project has to be implemented using Python Programming Language*.**
* ***Python Programming Language offers various features of implementing it*:-**
* **TTS :**– Google Text To Speech, for converting the given text to speech
* ***Speech recognition* :**– for recognizing the voice command and converting to text
* ***Selenium*: –** for web based work from browser
* ***Wolframalpha*:**– for calculation given by user
* ***Playsound*: –** for playing the saved audio file.
* ***Pyaudio*: –** for voice engine in python and many more as per requirements and further development of project.

**REFERENCE:**

1. **DETAILED STUDY OF PYTHON PROGRAMMING LANGUAGE.**

***https://***[***www.python.org***](http://www.python.org)***/***

***https://www.youtube .com/***

1. **DATA SCIENCE CONEPTS USING PYTHON**

[***https://www.edureka.co/***](https://www.edureka.co/)

***https://www.youtube.com/***

1. **MACHINE LEARNING CONCEPTS USING PYTHON**

[***https://www.stackoverflow.com/***](https://www.stackoverflow.com/)

***https://www.tutorialspoint.com/***

1. **NATURAL LANGUAGE PROCESS CONCEPTS USING PYTHON**

***https://www.youtube.com/***