CHAPTER 2

INTRODUCTION

2.1 Introduction to Voice Assistant:-

Many years ago, software programs were developed and run on the computer. Nowadays, computers are widely used by all people. Computers holds 44.59% of total market share worldwide. This shows that the market is increasing fast and there are also more capabilities for computer because of this wide use. Therefore, the software development on computer is very promising. It is not a convenient way for users with completely manually input. The common way of communication used by people in daily life is through the speech. If the computer can listen to the user for the request or handle the daily affairs, then give the right response, it will be much easier for users to communicate with their computer, and the computer will be much more "Smart" as a human assistant.

Fig 2.1 Worldwide market share of desktop vs mobile vs tablet

This project is focusing on the Desktop Assistant over the voice control (recognition, generate and analyse corresponding commands, intelligent responses automatically),

Google search and Google news, Wikipedia and input ranging from Speech-To-Text, Text-To-Speech technology.

Speech recognition has a long history with several waves of major innovations. Speech recognition for dictation, search, and voice commands has become a standard feature on smartphones and wearable devices. This project is a Personal Assistant with Voice Recognition Intelligence, which takes the user input in form of voice or text and process it and returns the output in various forms like action to be performed or the search result is dictated to the end user.

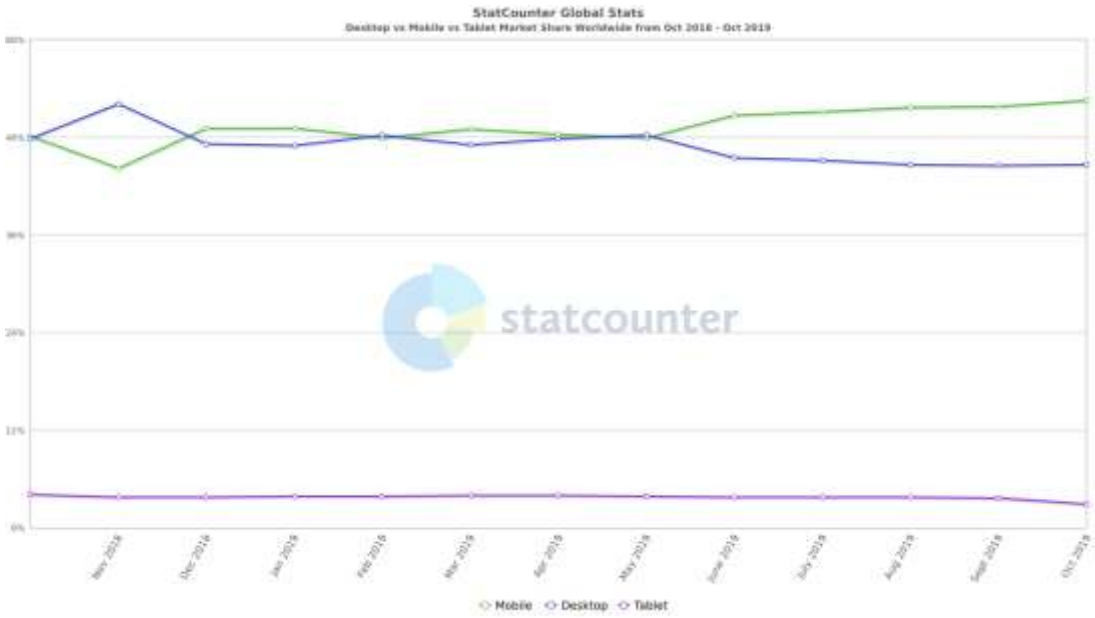

Fig 2.2 Graphical representation of worldwide market share of desktop vs mobile vs tablet

Design of a compact large vocabulary speech recognition system that can run efficiently on computer, accurately and with low latency. This is achieved by using the automated speech recognition (ASR) and Search components perform speech recognition and search tasks. In addition to ASR and Search, we also integrate a query parsing module between ASR and Search from a number of queries. Voice search is implemented as a two stage search procedure where string candidates generated by an automatic speech recognition (ASR) system are re-scored in order to identify the best matching entry from a number of specified queries.

2.2 About GARP:-

GARP (Gaurav Abhishek Ranjan Priya) has various branches of the services, but the main feature of GARP is Voice Recognition Engine which has an ability to work without internet connection i.e. Offline Voice Recognition. But anything which needs to be searched online, will surely require active internet connection.

The technology on which it is based upon is Python and Machine Learning. It performs its various tasks with the help of different python modules being imported in it for its functionality. It also uses the sub-modules of the main python modules. It performs its voice speaking ability with the help of inbuilt Microsoft voice package **SAPI5**. The **Speech Application Programming Interface** or **SAPI** is an API developed by Microsoft to allow the use of speech recognition and speech synthesis within windows application. To date, a number of versions of the API have been released, which have shipped either as part of a speech SDK or as part of Windows OS itself. Application that use SAPI includes Microsoft Office, Microsoft Agent and Microsoft Speech Server. It can operate in both male and female voice which can be accordingly set by their voice id. Voice id [0] is used for a male voice is which is called Microsoft Davis. Voice id [1] is used for a female voice which is called Microsoft Zira.

It has a function, named as takeCommand(), which takes microphone input from the user and convert it into string output. Subsequently, the output either is printed as a string on the output terminal or is converted into voice output with the help of pyttsx3 module and the assistant produces the speech output.

It gets turned on without any internet connection. Following is the list of things which it can perform based upon the verbal command being given to it:-

Task performed with internet:-

- Produces any search results from Wikipedia
- Sends Email.
- Opens YouTube.
- Opens Google.
- Googles about anything.
- Opens various shopping websites (Amazon, Flipkart, Myntra, Snapdeal, Shopclues, Jabong, Ajio)

- Opens various social media websites (Facebook, Instagram, Linked In, WhatsApp, Twitter)
- Shows us various news from google.

Task performed without internet:-

- Greets the user when being turned on.
- Plays music
- Plays videos
- Tells us the current time
- Starts any application which is installed in the system
- Terminates by itself by voice command of user

The various python modules used in GARP are:-

- pyttsx3
- PyAudio
- SpeechRecognition
- webbrowser
- datetime
- wikipedia
- os
- smtplib
- googlesearch
- urlopen
- BeautifulSoup