J.A.R.V.I.S: YOUR PERSONAL VIRTUAL ASSISTANT

MINOR PROJECT REPORT

SUBMITTED IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR

THE AWARD OF THE DEGREE OF

BACHELOR OF TECHNOLOGY

Information Technology

Submitted By:	Submitted To:
Japesh Bagga(1905423)	Prof. Harjot Kaur Gill
Daman(1805626)	Assistant Professor
Simple(1905430)	Minor Project Coordinator

Department of Information Technology

Guru Nanak Dev Engineering College,

Ludhiana-141006

Abstract

The project aims to develop a personal-assistant for Window-based systems. Jarvis draws its inspiration from virtual assistants like Cortana for Windows, and Siri for iOS. It has been designed to provide a user-friendly interface for carrying out a variety of tasks by employing certain well-defined commands. Users can interact with the assistant through voice commands. It first recognizes the user using face recognition and then proceeds with further process. As a personal assistant, Jarvis assists the end-user with day-to-day activities like searching queries in google, searching for videos, retrieving images, live weather conditions, sending emails, reporting date and time, word or any topic meanings directly from Wikipedia, tracking price of product, fetching COVID-19 Vaccination Drives facilities based on district, state or precisely by updating your pin code.

ACKNOWLEDGEMENT

We are highly grateful to Dr. Sehijpal Singh Principal ,Guru Nanak Dev Engineering College (GNDEC) Ludhiana, for providing this opportunity to carry out the minor project work on J.A.R.V.I.S Assistant.

The constant guidance and encouragement received from Dr. Kulvinder Singh Mann H.O.D IT Department, GNDEC Ludhiana has been of great help in carrying out the project work and acknowledgement with reverential thanks.

We, would like to express a deep sense of gratitude and thanks profusely to Prof. Hanit Karwal, without her wise counsel and able guidance, it would have been impossible to complete the project in this manner.

We express gratitude to other faculty of the Information Technology department of GNDEC for their intellectual support through out the course of this work.

Finally, We are indebted to all whosoever have contributed in this report work.

Japesh Bagga
Simple
Daman

Contents

1 Introduction

1.1 Introduction to Project Page

We are all admirers of the Marvel series and especially Tony Stark assistant J.A.R.V.I.S i.e. (Just A Rather Very Intelligent System) It is also very appropriate and fit for all he does for Tony. We have all dreamt of having one JARVIS for our whole lives. So, why not start by building one elementary JARVIS ourselves. It is a software agent that can perform tasks or services for an individual based on commands or questions. Though the real JARVIS is built with AI technologies, here in this project, we will code it with Python and already trained models for commonly used languages like English. All these models will be installed or imported as packages in our project.

Have you ever wondered how cool it would be to have your own A.I. assistant? Imagine how easier it would be to send emails without typing a single word, doing Wikipedia searches without opening web browsers, and performing many other daily tasks like playing music, opening text editors and other system softwareâs, tracking the price and comparing it with our budget, tracking all the available or ongoing COVID-19 vaccination drives with the help of a single voice command.

1.2 Objectives

1. To be capable enough to recognize the user using face recognition and then listen & respond to your voice.
2. To open most of the system software's like notepad, calculator, IDE's, Music Players, etc for you.
3. To be capable of opening websites like Google, YouTube, etc., in a web browser.
4. To be able to scrape some data from website like Wikipedia, E-commerce website and Co win.
5. To be able to track prices of desired products from e-commerce websites.
6. To do Web Scraping and report current date and time for you without even typing.
7. To send Email to the intended user without even typing the message and email address of the sender.
8. To be capable enough to fetch nearby COVID-19 Vaccination Drives according to the district, state or precisely by your pin code.

1.3 Unique Features of the System

1. You can track prices of the given product from the E-Commerce Websites according to the URL mentioned and can also tell it will be in your budget or not according to the budget set in settings.json file of the application.

2. You can track COVID-19 Vaccination Drives near your location according to the district, state or precisely by updating your pin code in the setting.json file.

3. It can work offline depending upon you are not requesting Jarvis to search on web or fetch anything from web.

1.4 Feasibility Study (Technical, Economical, Operational)

An AI personal assistant is a piece of software that understands verbal or written commands and completes task assigned by the client. It is an example of weak AI that is it can only execute and perform quest designed by the user. Speech recognition technology has become an increasingly popular concept in recent years. From organizations to individuals, the technology is widely used for various advantages it provides.

One of the most notable advantages of speech recognition technology includes the dictation ability it provides. With the help of technology, users can easily control devices and create documents by speaking. Speech recognition allows documents to be created faster because the software generally produces words as quickly as they uttered, which is usually much faster than a person can type. Dictation solutions are not only used by individuals but also by organizations that require massive transcription tasks such as healthcare and legal .

1. In technical aspect we have all the software and hardware required for this project. All the technologies required for this project like python libraries are easily accessed from the internet and are all loaded in the developerâs system.

2. In economical aspects all the softwares required for the project are easily accessible free of cost and legally from the internet.

3. In operational aspect , It is practically possible to develop this project and is operationally feasible.

2 System Design

2.1 Design Approach

We have used **Function Oriented Design approach** to software design where the design is decomposed into a set of interacting units where each unit has a clearly defined function. We have used various functions in each module to do some specific task. Some of them are as follows:-

Wish Me function():- In our project we make use of this wish me function() so that when the user start or run our program the Jarvis wish or greet the user according to the current time. we make you of this functionality by date & time module that helps the Jarvis to wish for current time.

Speak function() :- The first thing is that our AI assistant Jarvis should able to speak we make a speak function() that able to speak our Jarvis. speak function takes an audio argument and it will pronounce it we need to supply an audio to our speak function for that we install a pyttsx3 module with the pip installer and then import it . basically pyttsx3 is a python library function which covert text to speech. for voices we should use for the micro soft API sapi5 which provides two additional voices male and female.

Take Command function():- before using the take command function we need to install speech recognition or sr module and import it .the take command function taking return a string output by taking microphone input from the user.

Send mail function():- we make a send mail function with the help of this function our jarvis should send mail from another person. by the use of this email function firstly we need to import SMTP (simple mail transfer protocol) module this module allows us to send emails and to provide root for the different mails server.

2.2 System Design using Flowchart

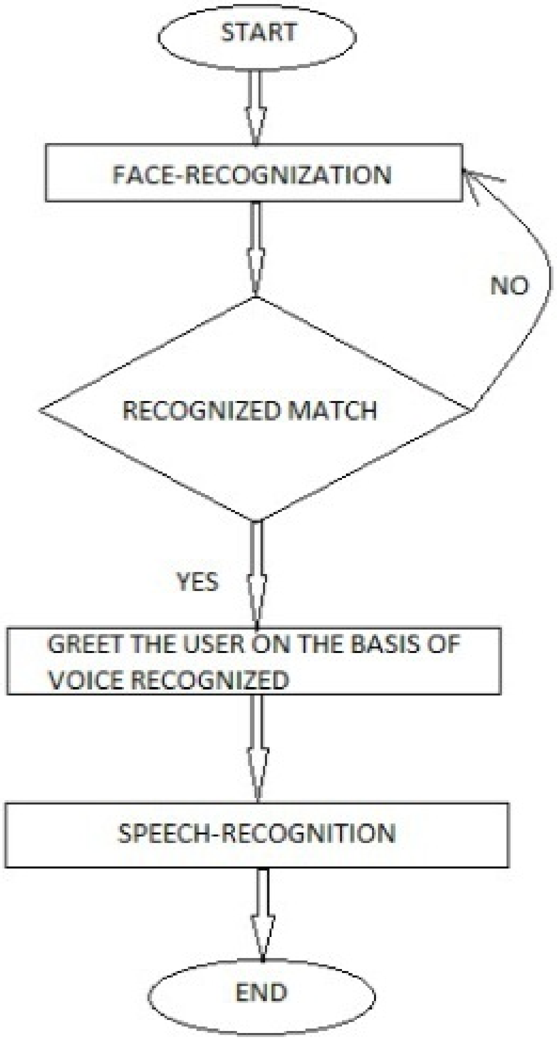


Figure 1: Project Flowchart

2.3 Methodology

Generally, it took lots of time to write code from scratch to build a Virtual Assistant. So, We have used various libraries and API's, which gives us easy functionality to build our Virtual personal Assistant JARVIS and add face recognition based security mechanism.

1. Recognizing & synthesis the voice
2. Training the speech recognition model to be able to understand different voices of people
3. Interpreting it using speech-to-text API
4. Processing the information, either by fetching the information from web using internet or offline only.
5. Sending the status or the result of the output in audio using text-to-speech functionality to user.

3 Implementation, Testing, and Maintenance

3.1 Introduction to Languages, IDE's, Tools and Technologies used for Implementation

3.1.1 Introduction to Languages

Python is an interpreted high-level general-purpose programming language. Python's design philosophy emphasizes code readability with its notable use of significant indentation. Its language constructs as well as its object-oriented approach aim to help programmers write clear, logical code for small and large-scale projects. It supports multiple programming paradigms, including structured (particularly, procedural), object-oriented and functional programming. Python is often described as a "batteries included" language due to its comprehensive standard library

Python's design offers some support for functional programming in the Lisp tradition. It has filter,mapandreduce functions; list comprehensions, dictionaries, sets, and generator expressions.The standard library has two modules (itertools and functools) that implement functional tools.

Features of Python :-

1. It is easy to understand and includes less development time
2. It is free and open soucre
3. It is high level language
4. It is portable user can works on Linux

3.1.2 IDE's

Visual Studio Code is a source-code editor made by Microsoft for Windows, Linux and mac-OS. Features include support for debugging, syntax highlighting, intelligent code completion, snippets, code refactoring, and embedded Git. In Visual Studio code users can change the theme, keyboard shortcuts, preferences, and install extensions that add additional functionality as their requirement and choice .

Visual Studio Code is a source-code editor that can be used with a variety of programming languages, including Java, JavaScript, Go, Node.js, Python and C++.Visual Studio Code includes multiple extensions for FTP, allowing the software to be used as a free alternative for web development. Code can be synced between the editor and the server, without downloading any extra software.Visual Studio Code allows users to set the code page in which the active document is saved, the newline character, and the programming language of the active document. This allows it to be used on any platform, in any locale, and for any given programming language.

GitHub is a provider of Internet hosting for software development and version control using Git. It offers the distributed version control and source code management functionality of Git, plus its own features. It provides access control and several collaboration features such as bug tracking, feature requests, task management, continuous integration and wikis for every project. GitHub offers its basic services free of charge. Its more advanced professional and enterprise services are commercial.[6] Free GitHub accounts are commonly used to host open-source projects.[we have also collaborate for GitHub in our project time and use some soucre code files at the implementation time.

3.1.3 Tools and Technologies used for Implementation

- We have used **“‘Co-WIN Public API“‘** to fetch all the nearby vaccination drive locations and details. Co-WIN Public APIs allow any third-party application to access certain un-restricted information, that can be shared with its users. This is limited only to read access in Co-WIN. The extent of access will be limited and in case of any misuse impacting the performance of Co-WIN solution will result in blocking any such application and entities as per the policies of MoHFW and taking any other appropriate action in accordance with law. The appointment availability data is cached and may be up to 5 minutes old. Further, these APIs are subject to a rate limit of 100 API calls per 5 minutes per IP. Please consider these points while using the APIs in your application.
- We have used **“‘Speech Recognition“‘** Library that is used for performing speech recognition, with support for several engines and APIs, both online and offline.
- To make our J.A.R.V.I.S. talk, we will make a function called speak(). This function will take audio as an argument, and then it will pronounce it. We must supply audio so that we can pronounce it using the speak() function we made. For this We are using a module called **“‘pyttsx3“‘** It’s a python library that will help us to convert text to speech. In short, it is a text-to-speech library. It works offline, and it is compatible with Python 2 as well as Python 3. In this module we have various TTS engines namely

- sapi5

- nsss

- espeak.

We have used **"'sapi5"'** application programming interface (API). It dramatically reduces the code overhead required for an application to use speech recognition and text-to-speech, making speech technology more accessible and robust for a wide range of applications. It Synthesize text strings and files into spoken audio using synthetic voices. Speech recognizers convert human spoken audio into readable text strings and files. Applications that use SAPI include Microsoft Office, Microsoft Agent and Microsoft Speech Server. In general, all versions of the API have been designed such that a software developer can write an application to perform speech recognition and synthesis by using a standard set of interfaces, accessible from a variety of programming languages. In addition, it is possible for a 3rd-party company to produce their own Speech Recognition and Text-To-Speech engines or adapt existing engines to work with SAPI. In principle, as long as these engines conform to the defined interfaces they can be used instead of the Microsoft-supplied engines.

In general, the Speech API is a freely redistributable component which can be shipped with any Windows application that wishes to use speech technology. Many versions (although not all) of the speech recognition and synthesis engines are also freely redistributable.

- To Recognize and manipulate faces from Python or from the command line with the worldâs simplest face recognition library **"'Face Recognition"'** built using dlibâs state-of-the-art face recognition built with deep learning. The model has an accuracy of 99.38you do face recognition on a folder of images from the command line! Dlib is a modern C++ toolkit containing machine learning algorithms and tools for creating complex software in C++ to solve real world problems. It is used in both industry and academia in a wide range of domains including robotics, embedded devices, mobile phones, and large high performance computing environments. Dlib's open source licensing allows you to use it in any application, free of charge. It internally uses OpenCV and Numpy for performing Image recognition or Processing.

- For Web Scraping, we have used **"'Beautiful Soup"'** It is a library that makes it easy to scrape information from web pages. It sits atop an HTML or XML parser, providing Pythonic idioms for iterating, searching, and modifying the parse tree.

4 Results and Discussions

4.1 Brief Description of Various Modules of the system

We will divide our whole project into 4 modules, that will be capable enough to work individually.

1. Face-Recognizer: It performs all the image recognition and processing tasks and outputs the name of the person of which the face is matched. The face detection and recognition is based on training data set that you can upload with your name in the Training set folder, from there only our module will be able to recognize the face of the user and will give output of the name.

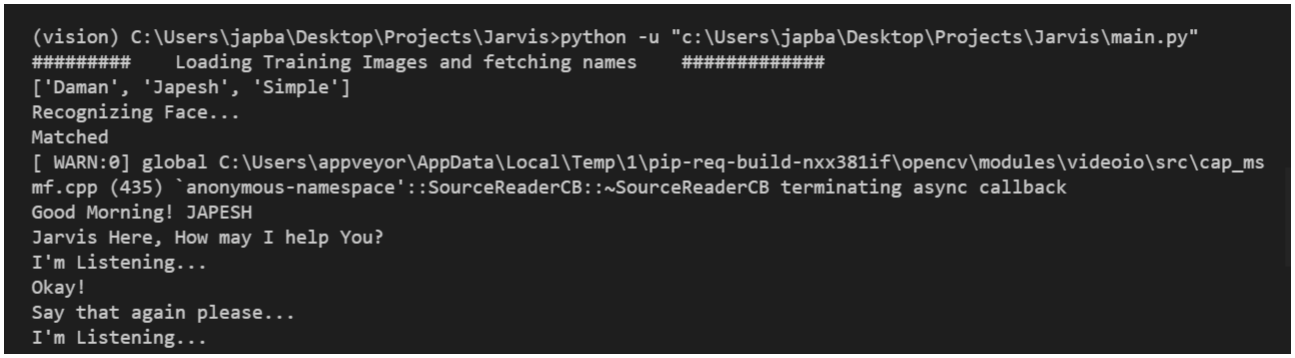
2. Jarvis(Speech-Recognizer) After getting the name of the user being detected, It speaks the name of him or her and then starts listening the voice of the user.

3. Price-Tracker: This module is being used inside Jarvis module to track the prices of the product we want to track then it's being compared with the budget we have set. According to the budget, it will tell that product is in our budget or not? You can set your product link and budget in the setting.json file.

4. Vaccine-Tracker: This module is also being used inside Jarvis module to track the COVID-19 Vaccination Drives near your location according to the district, state or precisely by updating your pin code in the setting.json file.

4.2 Snapshots of system with brief details

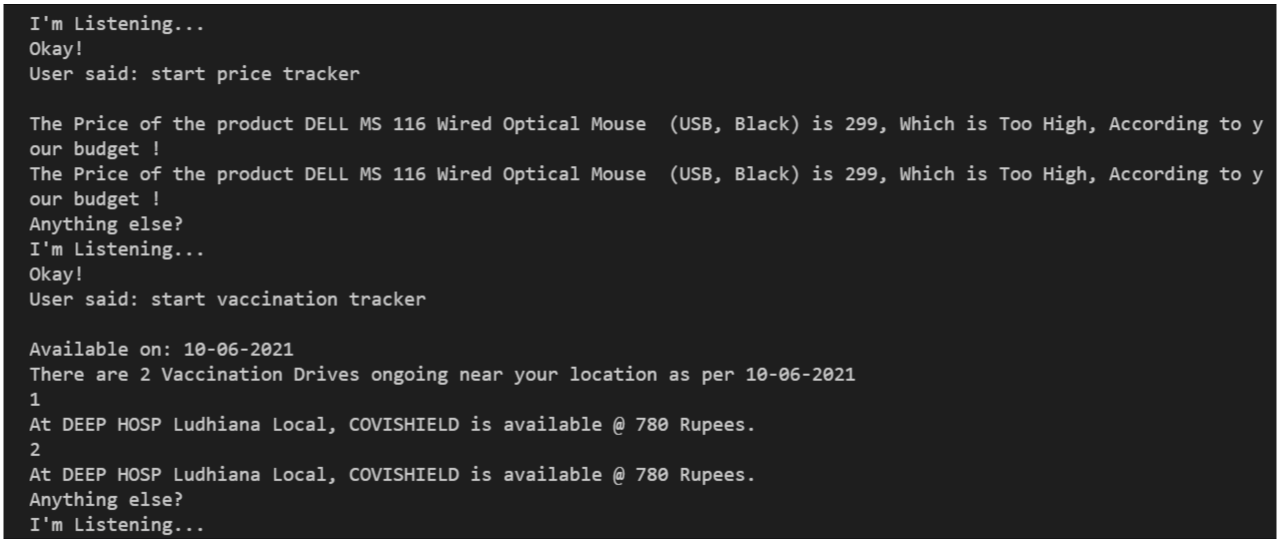
Though It's an not having any UI as it's more based on Speech and Audio but then also to show what is the flow of the program, we have printed all those sentences that our Jarvis and we will speak.



```
(vision) C:\Users\japba\Desktop\Projects\Jarvis>python -u "c:\Users\japba\Desktop\Projects\Jarvis\main.py"
#########    Loading Training Images and fetching names    #############
['Daman', 'Japesh', 'Simple']
Recognizing Face...
Matched
[ WARN:0] global C:\Users\appveyor\AppData\Local\Temp\1\pip-req-build-nxx381if\opencv\modules\videoio\src\cap_ms
mf.cpp (435) `anonymous-namespace'::SourceReaderCB::~SourceReaderCB terminating async callback
Good Morning! JAPESH
Jarvis Here, How may I help You?
I'm Listening...
Okay!
Say that again please...
I'm Listening...
```

Figure 2: Starter Program, Training Face Recognition Model & Recognizing After Getting Trained

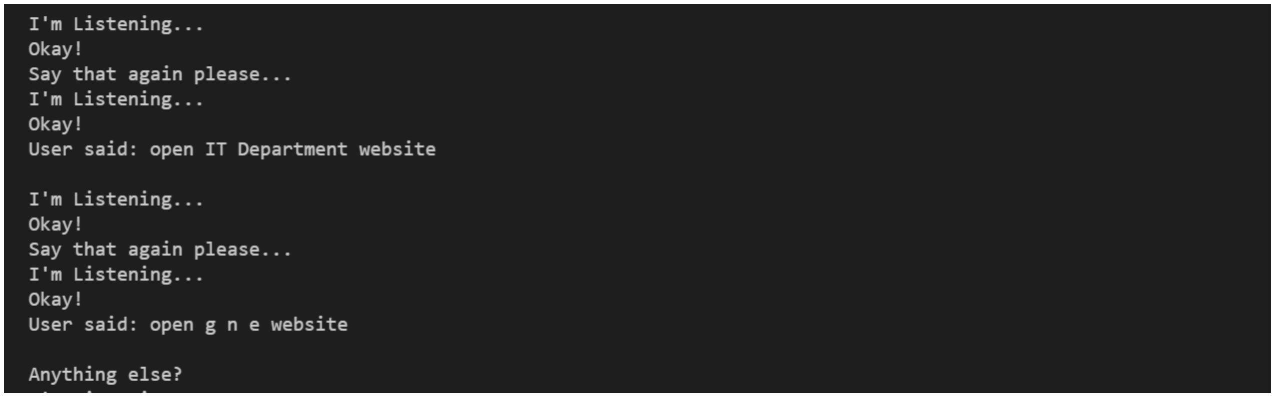
Firstly we have save some trained images in our project and we use them at the running time .When the user start running the project the face recognition security scan the user image and check it with user save trained image if the image is matched, then the Jarvis wish you according to the name mentioned in the image with which user has been matched. wishme() function is been used in our jarvis module that is working behind it.



```
I'm Listening...
Okay!
User said: start price tracker

The Price of the product DELL MS 116 Wired Optical Mouse  (USB, Black) is 299, Which is Too High, According to y
our budget !
The Price of the product DELL MS 116 Wired Optical Mouse  (USB, Black) is 299, Which is Too High, According to y
our budget !
Anything else?
I'm Listening...
Okay!
User said: start vaccination tracker

Available on: 10-06-2021
There are 2 Vaccination Drives ongoing near your location as per 10-06-2021
1
At DEEP HOSP Ludhiana Local, COVISHIELD is available @ 780 Rupees.
2
At DEEP HOSP Ludhiana Local, COVISHIELD is available @ 780 Rupees.
Anything else?
I'm Listening...
```

Figure 3: Price Tracker & Vaccination Drive Tracker Fetching Real Time Data from Web & Presenting

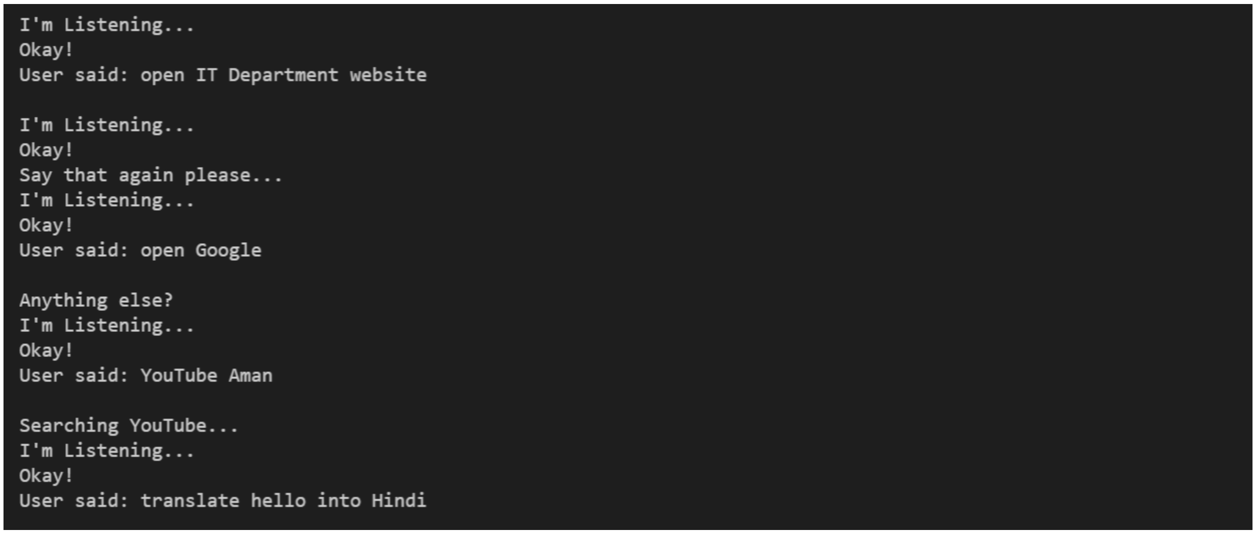
When the user gives the command ”start PRICE TRACKER ” to the Jarvis, it will fetch the data from online e-commerce shopping sites, whichever you have mentioned like Amazon, flipkart, then it will report you the price of product and will evaluate that is it under your budget, if yes it will be tell you that you can buy, it in your budge otherwise it will speak its price and will mention name. for eg. optical mouse if the price of optical mouse will under your budget then it will alert you with budget price and else it will alert the price is too high.



```
I'm Listening...
Okay!
Say that again please...
I'm Listening...
Okay!
User said: open IT Department website

I'm Listening...
Okay!
Say that again please...
I'm Listening...
Okay!
User said: open g n e website

Anything else?
```

Figure 4: Opening IT Department & GNDEC Website

If the user gives a command as ”open IT Department website”, Jarvis will open the IT Department website in the browser, without even typing a single letter, you will be able to see whole website.

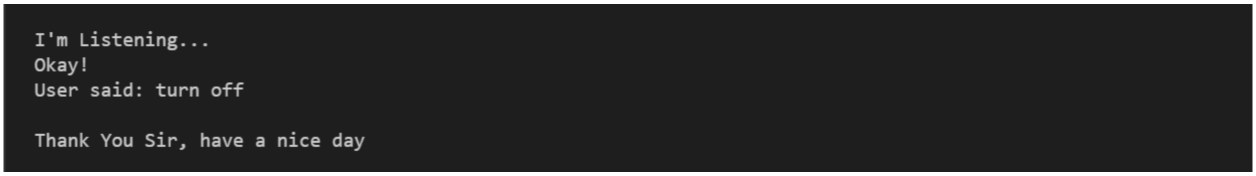


```
I'm Listening...
Okay!
User said: open IT Department website

I'm Listening...
Okay!
Say that again please...
I'm Listening...
Okay!
User said: open Google

Anything else?
I'm Listening...
Okay!
User said: YouTube Aman

Searching YouTube...
I'm Listening...
Okay!
User said: translate hello into Hindi
```

Figure 5: Opening Google Website, Searching Aman on YouTube & Translating hello into hindi

Here , we are using the elif loop to check whether Google is in the user’s query.If the user gives a command as ”open google” So, open google will be in the user’s query, and the elif condition will be true. Then in this case new browser tab will open with google.com opened.

Whenever, user gives a command as ”YouTube *sentence to be searched on YouTube*” Then it will search directly on YouTube the sentence told by user. we are using the elif condition loop to check whether user’s query is being start with YouTube or not.



```
I'm Listening...
Okay!
User said: turn off

Thank You Sir, have a nice day
```

Figure 6: Turning off the Jarvis

If the user gives a command as ”turn off” then the Jarvis will give a reply back ”Thank You Sir, have a nice day” and will get turn off which means it will not listen your command again.

5 Conclusion and Future Scope

Through this voice assistant, we have automated various services using a single line command. It eases most of the tasks of the user like searching the web, retrieving weather forecast details, vocabulary help and medical related queries. We aim to make this project a complete server assistant and make it smart enough to act as a replacement for a general server administration. The future plans include integrating Jarvis with mobile using React Native to provide a synchronised experience between the two connected devices. Further, in the long run, Jarvis is planned to feature auto deployment supporting elastic beanstalk, backup files, and all operations which a general Server Administrator does. The functionality would be seamless enough to replace the Server Administrator with Jarvis. We can also integrate other device and can make fully home or office automation system and notification facility either through SMS or Web Push notification.

References/Bibliography

[1] https://pyttsx3.readthedocs.io/en/latest/

[2] https://apisetu.gov.in/about.php

[3] https://docs.microsoft.com/en-us/previous-versions/windows/desktop/ms723627(v=vs.85)

[4] https://en.wikipedia.org/wiki/Microsoft$_{S}peech_{A}PI$

[5]$http://dlib.net/$

[6]$https://www.crummy.com/software/BeautifulSoup/bs4/doc/$