

## VIRTUAL ASSISTANT FOR WINDOWS USING PYTHON

#### **Elchuri Yaswanth Babu**

MCA Qis college of engineering and technology elchuriyaswanth2@gmail.com

#### **ABSTRACT**

In this modern era, day to day life became smarter and interlinked with technology. We already know some voice assistance like Google, Siri, etc. Virtual Personal Assistant is one of the most successful results of Artificial Intelligence, which has given a new way for the human to have its work done from a machine. This project gives a brief survey on the methodologies and concepts used in making of a Virtual Personal Assistant and thereby going on to use it in different software applications. Speech Recognition Systems, also known as Automatic Speech Recognition, plays an important role in virtual assistants in order to help user have a conversation with the system. In this project, we are trying to makea Voice based personal assistant which will include the important features that could help in assisting ones' needs. As the world is mostly being automated, we have heard some of the Virtual Assistants for windows like Alexa, Cortana, etc. Amazon hasn't added any features so that Alexa can control or access Windows specifically. Developing a Voice- Based Personal Assistant for Windows OS with the below main features. Opening any applicationthat we installed in the system. Artificial Neural Networks (ANN) to build an interactive assistant.

Keywords: Desktop Assistant, Pesonal Assistant, Voice Assistant, Speech Recognition, Python, Machine Learning, Virtual Assistant, Text to Speech, Speech to Text, Language Processing, Voice Recognition, Artificial Intelligence, Virtual Assistant.

#### **INTRODUCTION**

Knowingly, or unknowingly, personal assistants have become an integral part of our lives these days. It is because of Katti Jaya Krishna

M.Tech Qis college of engineering and technology jayakrishna.k@qiscet.edu.in

all the features and ease of use they provide. Personal Assistants are also capable of automating some day-to-day tasks, so that a user can focus on what matters the most to them. Features like, making calls, writing messages, taking photographs, storing todos on the go, browsing internet etc., are offered by personal assistants. So. utilization of these features of a virtual assistant will save an individual a lot of time, and effort. It is important to focus more on what matters the most for an individual, whether it could be personal work, or professional work. People often spend more time on doing routine tasks, and they can be automated with these types of personal assistants. When someone works in an environment with which he/she is not familiar with, they oftenfind it difficult to locate applications that they need, like browser, any IDE or nay other software. Most of the time, they will end up wasting hours of time, searching for the application alone. This results in unnecessary time wastage. Therefore, a voice based personal assistant will help automatingthis process. User is expected just to give a voice command, and the assistant will take care of the rest.

There is an unbelievable change rather advancement in technology over the last few years.Believe it or not, in today's world you can interact with your machine.

What is interacting with a machine? Obviously giving it some input, but what if the input is not in the conventional way of typing, rather it is your own Voice. What if you are talking to the machine, giving it commands and wanting the machine to interact with you like your assistant? What if the machine is not giving you answers just by showing you the best results but also by advising you with a better alternative. An easy access to machine with voice commands is the revolutionary way of human- system interaction. To achieve this, we need to use speech to text API for understanding the input. Many companies like Google, Amazon and Apple are trying

## PROBLEM STATEMENT

In today's world using personal assistants in smart phone is being more common but in desktops there are assistants like Alexa, Cortana which can able to get remainders, get destinations available on internet and get weather updates but these cannot able to play music, open, send messages on WhatsApp and these assistants don't have face authentication. We are implementing a desktop voice based personal assistant which has the features of these Cortana and Alexa along with these features we are adding more features like switching tabs, playing music, opening applications.

# **OBJECTIVES**

This project aims at developing a personal assistant for Windows-based systems. The main purpose of the software is to perform the tasks of the user at certain commands, provided in speech format. It will ease most of the work of the user as a complete task can be done on a single command.Jarvis draws its inspiration from virtual assistants like Cortana for Windows to achieve this in generalized form. Isn't it amazing that you can set reminders by just saying remind me to. Or set alarm with wake me up at .. Understanding the importance of this we have decided to make a system that can be placed anywhere in vicinity and you can ask it to help you do anything for you just by speaking with it. In addition to this, you can also connect two such devices through WiFi andmake them communicated with each other in future. This device can be very handy for day to day use and it can help you function better by constantly giving you reminders and updates. Why would we need it? Because vour own voice is turning into a best input device than a conventionalenter key. and Siri for iOS. Users can interact with the

assistant through voice commands. Currently, the project aims to provide the Windows users with a Virtual Assistant that would not only aid in their daily routine tasks like searching the web and many others. It also help in automation of various activites.

## **Literature Review**

**TOPIC:** Interacting with Computers by Voice: Automatic Speech Recognition and Synthesis

## AUTHOR: DOUGLAS

O'SHAUGHNESSY, SENIOR MEMBER, IEEE

This paper examines how people communicate with computers using speech. Automatic speech recognition (ASR) transforms speech into text, while automatic speech synthesis [or text-to- speech (TTS)] performs the reverse task. ASR has largely developed based on speech coding theory, while simulating certain spectral analyses performed by the ear. Typically, a Fourier transform is employed, but following the auditory Bark scale and simplifying the spectral representation with a decorrelation into cepstral coefficients. Current ASR provides good accuracy and performance on limited practical tasks, but exploits only the most rudimentary knowledge about human production and perception phenomena. The popular mathematical model called the hidden Markov model (HMM) is examined; first-order HMMs are efficient but ignore long-range correlations in actual speech. Common language models use a time window of three successive words in their syntactic– semantic analysis.

## **TOPIC**: An analysis of machine

synthesis has not yet been considered. Therefore, in this paper, we focus on machine translation and speech synthesis, and report a subjective evaluation to analyze the impact of each component. The results of these analyses show that the naturalness and intelligibility of synthesized speech are strongly affected by the fluency of the translated sentences.

The EMIME project is developing personalized S2ST, such that the a user's speech input in one language is used to produce speech output in another language. Speech characteristics of the output speech are adapted to the input speech characteristics using cross-lingual speaker techniques. adaptation While personalization is an important area of **TOPIC**: On the track of Artificial Intelligence: Learning with Intelligent Personal Assistants

AUTHORS: Nil Goksel, Mehmet Emin Mutlu

In a technology dominated world, useful and timely information can be accessed quickly via Intelligent Personal Assistants (IPAs). By the use of these assistants built into mobile operating translation and speech synthesis in speech-to-speech translationsystem

**AUTHORS**: Kei Hashimoto1, Junichi Yamagishi2, William Byrne3, Simon King2, Keiichi TokudaThis paper provides an analysis of the impacts of machine translation and speech synthesis on speechto-speech translation systems. The speechto-speech translation system consists of three components: speech recognition, machine translation and speech synthesis. Many techniques forintegration of speech recognition and machine translation have been proposed. However, speech

research, this paper focuses on the impactof speech machine translation and the synthesis components on endto-end performance of an S2ST system. In order to understand the degree to which each affects performance. component we investigate integration methods. We first conducted a subjective evaluation divided into three sections: speech synthesis, machine translation, and speech-to-speech translation. Various translated sentences were evaluated by using N-best translated sentences output from the machine translation component. The individual impacts of the machine translation and the speech synthesis components are analyzed from the results of this subjective evaluation.

systems, daily electronic tasks of a user can be accomplished 24/7. Such tasks like dictation, taking getting turn-by-turn directions, vocalizing email messages, daily appointments, setting reminding reminders, responding any factual questions and invoking apps can be completed by IPAs such as Apple's Siri, Google Now and Cortana. Microsoft The mentioned assistants programmed within Artificial Intelligence (AI) do create an interaction



between human and computer through a language used in digital natural communication. In this regard, the overall purpose of this study is to examine the potential use of IPAs that use advanced cognitive computing technologies and Natural Language Processing (NLP) for learning. To achieve this purpose, the working system of IPAs is reviewed briefly within the scope of AI that has recently become smarter to predict, comprehend and carry outmulti-step and complex requests of users.

#### METHODOLOGYSystem Architecture-

The overall system design consists of following phases:

- (a) Data collection in the form of speech
- (b) Voice analysis and conversion to text
- (c) Data storage and processing
- (d) Generating speech from the processed text output



## Fig - System Architecture of Voice Controlled Personal Assistant

In first phase, the data is collected

#### **EXISTING SYSTEM**

The existing windows OS has the inbuilt Cortana Assistant or Alexa assistant has able to get the definitions and quick answers available on the internet, get in the form of speech and stored as an input for the next phase for processing. In second phase, the input voice is continuously processed and converted to text using STT[1]. In next phase the converted text is analysed and processed using Python Script to identify the response to be taken against the command. Finally once the response identified, output is generated from simple text to speech conversion using TTS[2].

The basic architecture of the personal assistant is shown in figure.

#### System Components-IoT Box

The IoT box is used to perform operations for connecting smart devices[6]. The datatransmission takes place over WiFi network;

Task modules used in IoT box are:

- Servo Motor for mechanical movement for objects such as curtains. Here we are using this module to show how the curtain opens and closes itself; controlled by voice.
- We are using Neo-Pixel ring as lights to show the change of state of light and controllingbrightness of light.
- LED display(acting as a screen) that displays the data.IoT box architecture is shown in figure.



#### Fig - IoT box architecture

weather samples and set remainderetc. The existing system is not able to control the respective machine and it cannot able to do the activities which are not able to have in the application. It only able to do the static features it is not well designed to have the features which ae using daily by the people. Specifically windows is not having any these features and the assistants that exists

these features and the assistants that exists now are unable do any of the latest things likesending whatsapp messages etc.

## **PROPOSED SYSTEM**

We are going to develop a customized personal assistant which has the following features..

- Tab Switching, Basic Volume Controls, Screenshot.
- Opens any application that is installed in our machine.
- ➢ Minimizing and Maximizing windows.
- Playing Videos on YouTube, Getting information from google, Wikipedia.

# **REQUIREMENT ANALYSIS FUNCTIONAL REQUIREMENTS** Voice Recognition Requirements

The system must able to have the mic so that it can able to recognize the voice thatthe user is Saying. **Audio System** 

The system must have the audio output (Speaker) so that we can able to know the what the system is saying to us.

# NON-FUNCTIONAL REQUIREMENTS Usability

The system should be easy to use. The user should reach to the persons those who had fines with one button press if possible and also the user can able to edit thedetails of the fines oncehe paid in one click button. The system should be user friendly

# Reliability

This project is developed with deep learning and feature engineering techniques. Soin this step there is no certain reliable percentage that is measurable. Also, the input taken from the connected camera will be used to compare with the trained model and measure Reliability

# SOFTWARE REQUIREMENT SPECIFICATIONS

- Python IDLE or Visual Studio or Jupyter notebook.
- Windows OS.
- Any Browser like Google Chrome or Firefox or Microsoft Edge.
- PyQt5 For Designing Desktop Application.

# Why PyQts

PyQt5 is one of the most used modules in building GUI apps in Python, and that's dueto its simplicityas you will see. Another great feature that encourages developers to use PyQt5 is the PyQt5 designer, which makes it so easy to develop complex GUI apps in a short time. You just drag your widgets to build your form.

## Why Python

Python offers concise and readable code. While complex algorithms and versatile workflows stand behind machine learning and AI, Python's simplicity allows developers to write reliable systems. Python syntax is simple, which makes it easierto build models for machine learning and deep learning.

#### Why Browser

Browser is used to do most of the tasks like sending whatsapp messages, playing youtube videos and can able to send the emails also.

## Why Visual Studio Visual Studio

Code is a free coding editor that helps you start coding quickly. Use it to code in any programming language, without switching editors. Visual Studio Code has support for



many languages, including Python, Java, C++, JavaScript, and more. As you code, Visual Studio Code gives you suggestions to complete lines of code and quick fixes

HARDWARE SPECIFICATIONS

Processor : minimum i3.

RAM : Much need a minimum of 4GB RAMHard Disk : 1TB Desktop with Mic, Speaker / Laptop.

Deep learning models generally needs high processors because they have hidden layers and models need to learn the patterns by number of Epochs (iterating through input and output). So models require high processor. It is recommended to use cloud resources because buying GPU is costly.

#### SYSTEM ARCHITECTURE

Architecture can be referred to as a flow diagram, from where the user enters down to he CPU of the server and the power cord connected to it. To be more precise, the technologies, methods, and how everything is arranged to form a complete product is what the architecture of a system refers to. In this project the architecture consists of user interface, the user face consists of the buttons Run & Exit. Once the user or owner clicks the run button the application will start. The application will be open to receive the commands from the user. It receives the commands using the system mic. After that receiving the command it converts it to the text format, after converting it to the test format based upon the type of the text whether it is task or normal reply it goes with the respective task or interactive reply.

for common mistakes. You can also use the debuggerin VS Code to stepthrough each line of code and understand what is happening.



# Fig: Architecture Diagram for voice Based Personal Assistant using Deep Learnimg

#### **UML DIAGRAMS**

UML stands for Unified Modeling Language. UML is a language for specifying, visualizing, and documenting the system. This is the step while developing anyproduct after analysis. The goal from this is to produce a model of the entities involved in the project which later need tobe built.



# *Fig - Classification of UML Diagrams* USECASE DIAGRAM

Use case diagrams give a graphic

overview of the actors involved in a system, differentfunctions needed by those actors, and how these different functions interact. It's a great starting point for any project discussion because it helps in easily identifying the main actors involved and the main processes of the system. Use case diagram consists of use cases, actors and shows the interaction between the use case and actors. The purpose is to show the interactions between the use case and represent actorand to the system requirements from the user's perspective.

Actor: An actor in UML specifies a role played by a user or any other system that interacts with the subject.

#### Actors are:

#### Owner

**Use Case:** Use cases are a set of actions that some system or systems (subject) shouldor can performin collaboration with one or more external users of the system (actors).

#### Use Cases are:

- Application
- Take Command using Mic
- Converts Speech to Text
- Task Execution

#### **CLASS DIAGRAM**

Class diagrams contain icons representing classes, interfaces, and the relationships. You can create one or more class diagrams to represent the classes at the top level of the current model; such class diagrams are themselves contained by the top level of the current model. You can also create one or more class diagrams to represent classes contained by each package in your model such class diagrams are themselves contained by the package Actors are not part of the system. Actors represent anyone or anything that interacts with the system.Use case diagrams can be used during analysis to capture the system requirementsand to understand how the system should work. During the design phase, you can use use-case diagrams to specify the behavior of the system as implemented. Use case is asequence of transactions performed by a system that yields a measurable result of values for a particular actor.



Fig -Use Case Diagram For Voice Based Personal Assistant using Deep Learnin

enclosing the classes they represent the icons representing logical packages and classes in class diagrams.

- 1. Class diagrams are created to provide a picture or view of some or all of the casses in the model.
- 2. The main class diagram in the logical view of the model is typically a picture of thepackages in the' system. Each package also has its own main class diagram, which typically displays the "public" classes of the package.

A class diagram is a picture for

describing generic descriptions of possible systems. Class diagrams and collaboration diagrams are alternate representations of object models. A Class isa description of a group of objects with common properties (attributes) common behavior (operations), common relationships too their objects, and common semantics. Thus, a class is a template to create objects. Each object is aninstance of some class and objects cannot

#### Classes

User Interface : It has Run() and Exit() functions. If we click on the Run() function the application will start taking commands from the user. And the if we click on the Exit() function the application will be closed.

User : The user can interact with the Assistant and can give the commands to receive the responses from the assistant.

Take Command : Through the Mic the application takes the user's command as input.

Speaker : The Assistant gives the response to the user's command or input through the Speaker.

#### **ACTIVITY DIAGRAM**

An Activity diagram is a variation of a special case of a state machine, in which the states are activities representing the performance of operations and the transitions are triggered by the completion of the operations. The purpose of Activity diagram is to provide a view of flows and whatis going on inside a use case or among several classes. Activity diagrams contain activities, transitionsbetween the activities, decision points, and synchronization bars. Anactivity represents the performance of some behavior in the workflow. In the UML. activities are represented as rectangles with rounded edges, transitions

be instances of more than one class. In the UML, classes are represented as compartmentalized rectangles —

- 1. The top compartment contains the name of the class.
- 2. The middle compartment contains the structure of the class (attributes).
- 3. The bottom compartment contains the behavior of the class (operations).



## Fig - Class Diagram For Voice Based Personal Assistant using Deep Learning

are drawn as directed arrows, decision points are shown as diamonds, and synchronization bars are drawn as thick horizontal or vertical bars as shown in the following. The activity icon appears as a rectangle with roundedends with a name and a component for actions.





# Fig - Activity Diagram for Voice Based Personal Assistance using Deep Learning

#### SEQUENCE DIAGRAM

A sequence diagram is an interaction diagram that shows how processes operate with one another and in what order. It is a construct of a Message Sequence Chart. A sequence diagram shows object interactions arranged in time sequence. It depicts the objects and classes involved in the scenario and the sequence of messages exchanged between the objects needed to carryout the functionality of the scenario. Sequence diagrams are typically associated with use case realizations in the Logical View of the system under development. Sequence diagrams are sometimes called as event diagrams.



Fig - Sequence Diagram for Voice Based Personal Assistant using Deep Learning

# IMPLEMENTATION TECHNOLOGY DESCRIPTION KERAS

Keras is an open-source software library that provides a Python interface for artificial neural networks. Keras acts as an interface for the TensorFlow library. Up until version 2.3, Keras supported multiple backends, including TensorFlow, Microsoft Cognitive Toolkit, Theano, and PlaidML. As of version 2.4, only TensorFlow is supported. Designed to enable fast experimentation with deep neural networks, it focuses on being user-friendly, modular, and extensible.

It was developed as part of the research effort of project ONEIROS (Open-ended Neuro- Electronic Intelligent Robot Operating System), and its primary author and maintainer is Francois Chollet, a Google engineer. Chollet is also the author of the XCeption deep neural network model.

Keras contains numerous implementations of commonly used neuralnetwork building blocks such as layers, objectives, activation functions, optimizers, and a host of tools to make working with image and text data easier to simplify the coding necessary for writing deep neural network code. Thecode is hosted on GitHub, and community support forums include the GitHub issues page, and aSlack channel.In addition to standard neural networks, Keras has support for convolutional and recurrent neural networks.

It supports other common utility layers like dropout, batch normalization, and pooling. Keras allows users to productize deep models on smartphones on the web.

Deep learning is one of the major subfield of machine learning framework. Machine learning is the study of design of algorithms, inspired from the model of human brain.Deep learning is becoming more RERF

popular in data science fields like robotics, artificial intelligence(AI), audio & video recognitionand image recognition.

Artificial neural network is the core of deep learning methodologies.Deep learning is supported by various libraries such as Theano, TensorFlow, Caffe, Mxnet etc., Keras is one of the most powerful and easy to use python library, which is built on top of popular deep learning libraries like TensorFlow, Theano, etc., for creating deep learning models.

Keras runs on top of open source machine libraries like TensorFlow, Theano or Cognitive Toolkit(CNTK). Theano is a python library used for fast numerical computation tasks. TensorFlow is the most famous symbolic math library used for creating neural networks and deep learning models. TensorFlow is very flexible and the primary benefit is distributed computing.

CNTK is deep learning framework developed by Microsoft. It uses libraries such as Python, C#, C++ or standalone machine learning toolkits. Theano and TensorFlow are very powerful libraries but difficult to understand for creating neural networks. Keras based is on minimal structure that provides a clean and easy way to create deep learning models based on TensorFlow or Theano. Keras is designed to quickly define deep learning models. Well, Keras is an optimal choice for learning applications.

#### **Features of Keras**

Keras leverages various optimization techniques to make high level neural network API easier andmore performant. It supports the following features —

- Consistent, simple and extensible API.
- Minimal structure easy to achieve the result without any frills.

- It supports multiple platforms and backends
- It is user friendly framework which runs on both CPU and GPU.
- Highly scalability of computation.

# **Benefits Of Keras**

Keras is highly powerful and dynamic framework and comes up with the following advantages

- Larger community support.
- Easy to test.
- Keras neural networks are written in Python which makes things simpler.
- Keras supports both convolution and recurrent networks.
- Deep learning models are discrete components, so that, you can combine intomany ways.

# ARTIFICIAL NEURAL NETWORKS

Artificial Neural Networks (ANN) are multi-layer fully-connected neural nets that look like the figure below. They consist of an input layer, multiple hidden layers, and an output layer. Every nodein one layer is connected to every other node in the next layer. We make the network deeper by increasing the number of hidden layers.



*Fig 6.1.2.1: ANN Layer Model* If we zoom in to one of the hidden or output nodes, what we will encounter is the figure below.





#### Fig - Output Nodes of ANN

A given node takes the weighted sum of its inputs, and passes it through a nonlinear activation function. This is the output of the node, which then becomes the input of another node in the next layer. The signal flows from left to right, and the final output calculated by performing this procedure for all the nodes.

Training this deep neural network means learning the weights associated with all the edges. The equation for a given node looks as follows. The weighted sum of its inputs passed through a non- linear activation function. It can be represented as a vector dot product, where n is thenumber of inputs for the node.

$$z = f(b + x \cdot w) = f\left(b + \sum_{i=1}^{n} x_i w_i\right)$$
$$x \in d_{1 \times n}, w \in d_{n \times 1}, b \in d_{1 \times 1}, z \in d_{1 \times 1}$$

#### Fig : Mathematical Representation for Neural Network

So far we have described the forward pass, meaning given an input and weights how the output is computed. After the training is complete, we only run the forward passto make the predictions. But we first need to train our model to actually learn the weights, and the training procedure works as follows:

Randomly initialize the weights for

all the nodes. There are smart initializationmethods which we will explore in another article. For every training example, perform a forward pass using the current weights, and calculate the output of each node going from left to right. The final output is the value of the last node. Compare the final output with the actual target in the training data, and measure the error using a loss function. Perform a backwards pass from right to left and propagate the error to every individual node using backpropagation. Calculate each weight's contribution to he error, and adjust the weights accordingly using gradient descent. Propagate the error gradients back starting from the last layer.

Neural architecture search (NAS)uses machine learning to automate ANN design. Various approaches to NAS have designed networks that compare well with hand-designed systems. The basic search algorithm is to propose a candidate model, evaluate it against a dataset and use the results as feedback to teach the NAS network.[ Available systems include AutoML and AutoKeras. Design issues include deciding the number, type and connectednessof network layers, as well as the size of each and the connection type (full, pooling,). Hyperparameters must also be defined as part of the design (they are not learned), governing matters such as how many neuron sare in each layer, learningrate, step, stride, depth, receptive field and padding (for CNNs), etc.

Using Artificial neural networks understanding requires an of their characteristics. Choice of model: This depends on the data representation and the application. Overly complex models are slow learning. Learning algorithm: Numerous trade-offs exist between learning algorithms. Almost any algorithm will work well with the correct hyperparameters for training on a particular data set. However, selecting and tuningan algorithm for training on unseen data requires significant experimentation.

Robustness: If the model, cost function and learning algorithm are selected appropriately, the resulting ANN can become robust.

ANN capabilities fall within the categories: following broad Function approximation, or regression analysis, including time series prediction, fitness and approximation modeling. Classification, including pattern and sequence recognition, noveltydetection and sequential decision making. Data processing, including filtering, clustering, blind source separation and compression. Robotics, including directing manipulators and prostheses. Applications of ANN:

Because of their ability to reproduce and model nonlinear processes, artificial neural networkshave found applications in many disciplines. Application areas includesystem identification and control (vehicle control, trajectory prediction, natural process control. resource management), quantum chemistry, general game playing, pattern recognition (radar systems, face identification, signal classification, 3D reconstruction, object recognition and more), sensor data analysis, sequence recognition (gesture, speech. handwritten and printed text recognition), medical diagnosis, finance (e.g. automated trading systems), data mining, visualization, machine translation, social network filtering and e-mail spam filtering. ANNs have been

#### **IMPORTED MODULES:**

used to diagnose several types of cancers and todistinguish highly invasive cancer cell lines from less invasive lines using only cell shape information.

ANNs have been used to accelerate reliability analysis of infrastructures subject to natural disasters and to predict foundation settlements. ANNs have also been used for building black-box models in geoscience. Building black-box models in geoscience: hydrology, ocean modelling and coastal engineering, and geomorphology. ANNs have been employed in cybersecurity, with the objective to discriminate between legitimate activities and malicious ones.

For example, machine learning has been used for classifying Android malware, for identifyingdomains belonging to threat actors and for detecting URLs posing asecurity risk. Research is underway on ANN systems designed for penetration testing,for detecting botnets, credit cards fraudsand network intrusions.

ANNs have been proposed as a tool to solve partial differential equations in physics and simulate the properties of many-body open quantum systems. In brain research ANNs have studied short-term behavior of individual neurons, the dynamics of neuralcircuitry arise from interactions between individual neurons and how behavior can arise from abstract neural modules that represent complete subsystems.

Studies considered long-and short-term plasticity of neural systems and their relation tolearning and memory from the individual neuron to the system level.



import pyttsx3 # for speaking import requests import speech\_recognition as sr import datetime import os import time import sys import webbrowser

#### **SPEECH RECOGNITION:**

The voice module used this system is Google's API i.e. "import speech\_ recognition as sr".

This module is used to recognize the sound waves given by the user as input.

This is a loose API this is supplied and supported by Google. This is a totally mild API that facilitates in decreasing the scale of our application.

#### TTS & STT:

#### DATETIME:

The Date-Time module is imported to support the date and time. For example, the consumer wants to recognize the modern- day date and time or the person desires to time table a venture at a sure time. In brief this module helps instructions to manipulate date and time and carry out operationsaccording to it handiest. This is a critical module,mainly in tasks in which we need to keep a track of time. This module could be very small in length and allowscontrolling the dimensions of our program. If the modules are too large or heavy then the system will lag and provide gradual responses.

#### **WEBBROWSER:**

Web-browser module is imported to display information from web to users. If

The input voice is first converted to text by using speech recognition module. The text is thenprocessed to result of the voice by the user. The most time ingesting a number of the STT because the gadget first has to concentrate to the user and unique users have distinctive, a few are smooth to apprehend whilst a few are not without difficulty audible. Once the speech is converted to text executing commands and giving the consequences lower back to the user isn't always a time-eating.

#### PYTTSX3:

To convert text into speech in python the pyttsx3 module is used. This is an offline module.

The module provides run and wait functionality. It is used to allow how much time the system will wait for another input of user. This is a module available in the python community for free that can be installed using the pip command.

the user wants to open browser and gives input as "Open Google". Then input is processed using this module and the Google browser is opened. The browserwhich is set in code will open.

#### **OS MODULE:**

OS Module provides operating system dependent functionalities. If we want to perform operations of OS like data reading, data writing, or data manipulate paths then this types of functions are available in an OS module. When the these operations raise an error like "OSError" in case of anyerror like invalid names, paths, or arguments which may be incorrect or correct but just not accepted by the operating system.

PROCEDURE FOR EXECUTIONStep-

1: Open Command Prompt







Fig: 6.2.2: Changed to the directory where our python is present

**Step-2**: Navigate or change the directory where your python file is located.



*Fig - Executing master python* **Step-3**: Execute the master python file using python filename.py command.

🔲 0 💌 🖿 / GEC / My work / Major Project
C
🗋 🗅 owner
🗋 🛢 jarvis.ipynb
C Contacts.vcf
□ □ JARVIS.gif
🗋 🗋 jarvis.mp3
🗋 🗅 jarvis.py
☐ ☐ jarvis_low_battery.mp3
🗋 🗅 jarvisUi.py
🗋 🗋 jarvisUi.ui
pywhatkit_dbs.txt
Welcome Home Sir.mp3

Fig - Project Structure

**Jarvis.ipynb:** This is the main master file that has all the source code in it.

**Jarvis.gif:** This the gif file, it is in integration with the desktop application that we have created.

**JarvisUi.py:** This file is used to have all the code related to our desktop applicationor it has allthe code related to our user interface.

**Mp3 files:** All Mp3 files are used for the welcome message or for the indications.

## CONCLUSION

Development of a Voice based personal assistant for desktop using Python programming language. This Voice based personal assistant, in today's life style will be more effective incase of saving time, compared to that of previous days. This Personal Assistant has been designed with ease of use as the main feature. The Assistant works properly to perform some tasks given by user. Furthermore, there are many things that this assistant is capable of doing like opening any application in the



system and playing music with just one voice command.

#### **FUTURE SCOPE**

This application can be used in various situations. We can enhance the project by adding the face Recognition feature. So that the application can recognize it's owner. We can enhance our projectby using CNN for the recognizing the face more accurately and also able to make the assistant to dothe other things like noting down the notes by opening any application and able to run the respective programming scripts with its respective compiler.

#### BIBLIOGRAPHY

- DOUGLAS O'SHAUGHNESSY, SENIOR MEMBER, IEEE, "Interacting With Computers by Voice: Automatic Speech Recognition and Synthesis" proceedingsof THE IEEE, VOL. 91, NO. 9, SEPTEMBER 2003
- 2. Kei Hashimoto1, Junichi Yamagishi2, William Byrne3, Simon King2, Keiichi Tokuda, "An analysis of machine translation and speech synthesis in speech-to- speech translation system" proceedings of5108978-1-4577-0539- 7/11/\$26.00 ©2011 IEEE.
- **3.** Nil Goksel-Canbek2 Mehmet Emin Mutlu, "On the track of Artificial Intelligence: Learning with Intelligent Personal Assistant" International Journal of Human Sciences, 13(1), 592-601. doi:10.14687/ijhs.v13i1.3549.
- **4.** Harshita Phatnani, Mr. Jyotiprakash Patra and Ankit Sharma' "CHATBOT ASSISTING: SIRI" Proceedings of BITCON-2015 Innovations For National Development National Conference on : Research and Development in Computer Science and Applications, E- ISSN2249–8974.
- Sutar Shekhar, Pophali Sameer, Kamad Neha, Prof. Devkate Laxman, " An Intelligent VoiceAssistant Using Android Platform", March 2015, IJARCSMS, ISSN: 232 7782
- 6. VINAY SAGAR, KUSUMA SM, "Home Automation Using Internet of Things", June- 2015, IRJET, e-ISSN: 2395-0056
- 7. <u>https://www.cnet.com/products/ivee-sleek/reviews/</u>
- 8. <u>https://www.cnet.com/products/athom-homey/preview/</u>