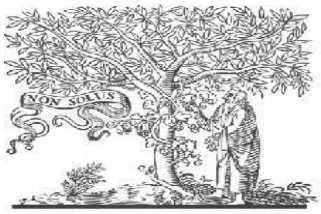
International Journal for Innovative Engineering and Management Research
A Peer Revieved Open Access International Journal
www.ijiemr.org

COPY RIGHT



ELSEVIER
SSRN

IJIEMR Transactions, online available on 07th Jan 2023. Link
:http://www.ijiemr.org/downloads.php?vol=Volume-12&issue=ISSUE-1

DOI: 10.48047/IJIEMR/V12/ISSUE 01/58

Title **SRIDEVI VOICE ASSISTANT**

Volume 12, Issue 1, Pages: 618-627

Paper Authors
Mr. R.Sreedhar, Ch. Anantha Kala, Beeram Indhu, Pulipati Kavitha, N.Bhavana, Rayini Ramya Sree

USE THIS BARCODE TO ACCESS YOUR ONLINE PAPER

To Secure Your Paper As Per UGC Guidelines We Are Providing A Electronic Bar Code

International Journal for Innovative Engineering and Management Research
A Peer Revieved Open Access International Journal
www.ijiemr.org

SRIDEVI VOICE ASSISTANT

Mr. R.Sreedhar, Associate Professor, Dept. of Information Technology, Sridevi Women's Engineering College, Hyd. rachasreedharswec@gmail.com

Ch. Anantha Kala, B.Tech., Dept. of Information Technology, Sridevi Women's Engineering College, Hyd.

Beeram Indhu, B.Tech., Dept. of Information Technology, Sridevi Women's Engineering College, Hyd.

Pulipati Kavitha, B.Tech., Dept. of Information Technology, Sridevi Women's Engineering College, Hyd.

N.Bhavana, B.Tech., Dept. of Information Technology, Sridevi Women's Engineering College, Hyd.

Rayini Ramya Sree, B.Tech., Dept. of Information Technology, Sridevi Women's Engineering College, Hyd.

ABSTRACT:

The project aims to develop a personal-assistant for Laptops. Sridevi draws its inspiration from virtual assistants like Cortana for Windows, and Siri for iOS but Cortana extends its services to some extent. Our project has been designed to provide a user-friendly interface for carrying out a variety of tasks by employing certain well-defined activities which made it easier to search on Google without opening the browser, and performing many other daily tasks like playing music, opening your command prompt with the help of a single voice command. Users can interact with the assistant either through voice commands or using keyboard inputs. Python is an emerging language so it becomes easy to write a script for Voice Assistant in Python. The instructions for the assistant can be handled as per the requirement of user. In Python there is an API called Speech Recognition which allows us to convert speech into text. It was an interesting task to make our own assistant. By making this project, we realized that the concept of AI in every field is decreasing human effort and saving time.

Keywords : Voice Assistant, Python, Machine Learning, Text-Speech-Text, Voice Detection and Artificial Intelligence.

INTRODUCTION

Artificial Intelligence when used with machines, it shows us the capability of thinking like humans. In this, a computer system is designed in such a way that typically requires interaction from human.

As we know Python is an emerging language so it becomes easy to write a script for Voice Assistant in Python. The instructions for the assistant can be handled as per the requirement of user. Speech recognition is the Alexa, Siri, etc. In Python there is an API called Speech Recognition which allows us to convert speech into text. It was an interesting task to make my own assistant. It became easier to send emails without typing any word, searching on Google without opening the browser, and performing many other daily tasks like playing music, opening your favorite IDE with the help of a single voice command. In the current scenario, advancement in technology is such that they can perform any task with same effectiveness or can say more effectively than us. By making this project, I realized that the concept of AI in every field is decreasing human effort and saving time.

The functionalities include, it can send emails, it can read PDF, It can send text on WhatsApp, It can open command prompt, your favorite IDE, notepad etc., It can play music, It can do Wikipedia searches for you, It can open websites like Google, YouTube, etc., in a web browser, It can give weather forecast, It can give desktop reminders of your choice. It can have some basic conversation.

Tools and technologies used are PyCharm IDE for making this project, and I created all py files in PyCharm. Along with this I used following modules and libraries in my project. pyttsx3, Speech Recognition, Datetime, Wikipedia, Smtplib, pywhatkit, pyjokes, pyPDF2, pyautogui, pyQt etc.

PURPOSE

Recently the usage of virtual assistants to control your surroundings is becoming a common practice. We make use of Google AI, Siri, Alexa, Cortana, and many other similar virtual assistants to complete tasks for us with a simple voice or audio command. You could ask them to play music or open a particular file or any other similar task, and they would perform such actions with ease.

While these devices are cool, it is also intriguing to develop your own AI voice automated assistant, which you can utilize to control your desktop with just the help of your voice. We can use such an AI to talk with you, open videos, play music, and so much more.

We will work on developing an introductory project for a Voice Assistant that you can utilize to control your PC or any other similar device with your voice. We will get started with an introduction to some

of the basic dependencies that are required to construct this project and proceed to put it all together into a Python file through which the AI Voice assistant is built to follow your commands.

SCOPE

The main objective of Voice Assistant is to make people smart and provide them with instant and accurate results. It takes the voice input and converts it into a computer-readable language that can answer their questions. The goal of this service is to provide users with results that they have asked for through the web.

LITERATURE SURVEY

This area of digital assistants having speech reputation has visible some primary advancements or inventions. This is especially due to its call for in gadgets like smart watches or health bands, speakers, Bluetooth earphones, cellular telephones, computer or desktop, TV, and so forth. Almost all the digital gadgets which are coming nowadays with voice assistants, which assist to control the device with speech recognition. A new set of strategies is being evolved constantly to improve the performance of voice computerized seek. With using voice assistants, we are able to automate the project without difficulty, simply give the center to the machine within the speech shape and all the duties might be accomplished by means of it from changing your speech into textual content shape to putting off keywords from that text and execute the question to provide outcomes to the person. This has been one of the most beneficial improvements in era. Before AI we have been the ones who have been upgrading technology to do a assignment however now the gadget is itself able to counter new responsibilities and clear up it without need to involve the people to conform it.

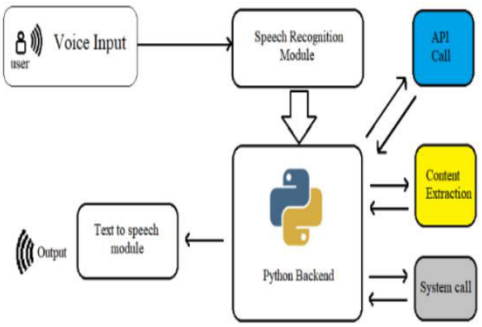


Fig no.1 Process of a System

This field of virtual assistants having speech recognition has seen some major advancements or innovations. This is mainly because of its demand in devices like smartwatches or fitness bands, speakers, Bluetooth earphones, mobile phones, laptop or desktop, television, etc. Almost all the digital devices which are coming nowadays are coming with voice assistants which help to control the device with speech recognition only.

A new set of techniques is being developed constantly to improve the performance of voice automated search. As the amount of data is increasing exponentially now known as Big Data the best way to improve the results of virtual assistants is to incorporate our assistants with machine learning and train our devices according to their uses. Other major techniques that are equally important are Artificial Intelligence, Internet Of Things, Big Data access and management, etc.

With the use of voice assistants, we can automate the task easily, just give the input to the machine in the speech form and all the tasks will be done by it from converting your speech into text form to taking out keywords from that text and execute the query to give results to the user. Machine Learning is just a subset of Artificial

Intelligence. This has been one of the most helpful advancements in technology. Before AI we were the ones who were upgrading technology to do a task but now the machine is itself able to counter new tasks and solve it without need to involve the humans to evolve it. This has been helpful in day-to-day lifestyle.

From mobile phones to personal desktops to mechanical industries these assistants are in very much demand for automating tasks and increasing efficiency.

An overview of Speech Recognition

Speech recognition is a technology that able a computer to capture words spoken by a human with a help of microphone. These words are later on recognized by speech recognizer, and in the end, system outputs the recognized words. The process of speech recognition consists of different steps that will be discussed in the following sections one by one.

An ideal situation in the process of speech recognition is that, a speech recognition engine recognizes all words uttered by a human but, practically the performance of a speech recognition engine depends on number of factors. Vocabularies, multiple users and noisy environment are the major factors that are counted in as the depending factors for a speech recognition engine. 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 The input voice is first converted to text by using speech recognition module. The text is then processed 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.

EXISTING SYSTEM:

We are still far away from understanding the genuine capability of speech recognition technology. This applies both to the refinement of the innovation itself and to its coordination into our lives. The current digital assistants can decipher discourse great, yet they are not the conversational interfaces that the innovation suppliers need them to be. Besides, speech recognition stays constrained to few items. Voice recognition technology was around some time before Apple's Siri appeared in 2011. At the Seattle World's fair in 1962, IBM presented a device called Shoebox. It was the size of a shoebox and could perform scientific functions and perceive 16 spoken words as well as digits 0-9. Mozilla are working on foundations for open, public voice services. Training a voice assistant takes a lot of data though: 10,000 hours of recordings. To put that in context, the total of all of the TED talks out there comes to about 100 hours: still 2 orders of magnitude away! That's why Mozilla have opened up Project Common Voice to allow the public to lend their own voices. They created an experiment called Voice Fill to allow you to search via voice on Google, Yahoo and duckduckgo. And they have been starting to explore the idea of open voice service registration under a preliminary name of Voice HTML.

The current voice assistant system basically existing on Windows OS is the Cortana which is completely online based system and requires high speed fast internet and also a regular Microsoft account for login and other existing system is Ok-Google voice assistant which is browser dependent.

Disadvantages:

- Lack of Accuracy and Misinterpretation
- Time Costs and Productivity

- Accent and speech recognition

3.1.2 problem statement

Artificial intelligence personal assistants have become plentiful over the last few years. applications such as Siri, Bixby, ok google and Cortana make mobile device users' daily routines that much easier. you may be asking yourself how these functions. well, the assistants receive external data (such as movement, voice, light, Gps readings, visually defined markers, etc.) via the hardware's sensors for further processing - and take it from there to function accordingly. not too long ago, building an ai assistant was a small component of developers' capacities; however, nowadays, it is quite a realistic objective even for novice programmers. to create a simple personal ai assistant, one simply needs dedicated software and around an hour of working time. it would take much more time, though, to create something more advanced and conceptually innovative. nonetheless, well thought-out concepts can result in a great base for a profitable startup. let us consider the six most renowned applications based on artificial intelligence concepts that can help create your virtual ai assistant app.

We are all well aware about Cortana, Siri, google assistant and many other virtual assistants which are designed to aid the tasks of users in windows, android and iOS platforms. but to our surprise, there's no such complete virtual assistant available for core windows platform consisting of 70% of the users. so, this is actually a major problem for users where there could be internet instability, server problems and places where internet is not accessible**.**

PROPOSED SYSTEM

The work is initialized with analyzing the audio commands given by the user via microphone. This can be anything like retrieving any information, operating computer's files, etc. Tests are conducted by programming according to books and online resources, with the goal to find best practices and a more advanced understanding of Voice Assistant. Fig.1 shows the detailed workflow of the basic process of the voice assistant.

Speech recognition is used to convert the speech input to text. This text is then fed to the central processor which determines the nature of the command and calls the relatable script for execution. But the difficulties don't end there. Even with tons of hours of input, other factors aside can play a big role in whether or not the software can understand you basically. Background noise can easily eliminate a speech recognition device off the track. This is because it does not inherently have the ability to classify the ambient sounds it "hears" of a dog barking or a helicopter flying overhead, from your voice. Developers have to program that ability into the machine; they conduct data collection of these ambient sounds and "tell" the device to filter them out accordingly.

Another factor is the way humans naturally shift the pitch of their voice to accommodate for noisy environments; speech recognition systems can be sensitive to these pitch changes in most of the conditions. In our system as specified the input process goes on till the end which is double-edged sword

as user on one side has ended his speech input but the recognizer waits for an ample time before it switches and converts the subsequent input to text. Then the Speech Recognition module now completes the recognizing process and generates a search thread.

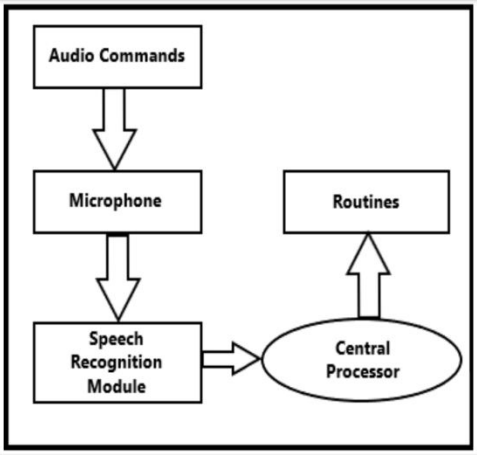


Fig . Proposed System

ADVANTAGE:

- Easily configuration to perform many of regular tasks by simply giving voice commands.
- Voice based search that is a boon for many like senior citizens who are not comfortable using the keypad/keyboard.
- Able to write the text through both keyboard and voice input.

Requires less consumption of time in writing text. Open different windows software's, based on voice input.

METHODLOGY

With the advancement SRIDEVI can perform any task with same effectiveness or can say more effectively than us. By making this project, I realized that the concept of AI in every field is decreasing human effort and saving time. Functionalities of this project include, It can send emails, It can read PDF, It can send text on WhatsApp, It can open command prompt, your favorite IDE, notepad etc., It can play music, It can do Wikipedia searches for you, It can open websites like Google, YouTube, etc., in a web browser, It can give weather forecast, It can give desktop reminders of your choice. It can have some basic conversation.

The data flow for SRIDEVI is as follow

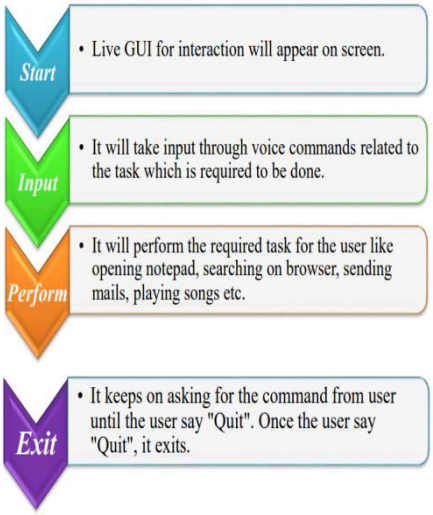


Fig no.. DATA FLOW

The system is designed using the concept of Artificial Intelligence and with the help of necessary packages of Python. Python provides many libraries and packages to perform the tasks, for example pyPDF2 can be used to read PDF. The details of these packages are mentioned in Chapter 3 of this report. The data in this project is nothing but user input, whatever the user says, the assistant performs the task accordingly. The user input is

nothing specific but the list of tasks which a user wants to get performed in human language i.e. English.

The IDE used in this project is PyCharm. All the python files were created in PyCharm and all the necessary packages were easily installable in this IDE. For this project following modules and libraries were used i.e. pyttsx3, SpeechRecognition, Datetime, Wikipedia, Smtplib, pywhatkit, pyjokes, pyPDF2, pyautogui, pyQt etc. I have created a live GUI for interacting with the JARVIS as it gives a design and interesting look while having the conversation

Speech Synthesis (or Text to Speech) is the computer-generated simulation of human speech. It converts human language text into human-like speech audio. In this Project, one will learn how you can convert text to speech in Python. we won't be building neural networks and training the model in order to achieve results, as it is pretty complex and hard to do it. Instead, we are going to use some APIs and engines that offer it. There are a lot of APIs out there that offer this service, one of the commonly used services is Google Text to Speech, we will play around with it along with another offline library called pyttsx3

7.1 S speech Recognition

Speech recognition involves receiving speech through a device's microphone, which is then checked by a speech recognition service against a list of grammar (basically, the vocabulary you want to have recognized in a particular app.) When a word or phrase is successfully recognized, it is returned as a result (or list of results) as a text string, and further actions can be initiated as a result.

TYPES OF TESTS

Unit testing

Unit testing involves the design of test cases that validate that the internal program logic is functioning properly, and that program inputs produce valid outputs. All decision branches and internal code flow should be validated. It is the testing of individual software units of the application .it is done after the completion of an individual unit before integration. This is a structural testing, that relies on knowledge of its construction and is invasive. Unit tests perform basic tests at component level and test a specific business process, application, and/or system configuration. Unit tests ensure that each unique path of a business process performs accurately to the documented specifications

and contains clearly defined inputs and expected results.

Integration testing

Integration tests are designed to test integrated software components to determine if they actually run as one program. Testing is event driven and is more concerned with the basic outcome of screens or fields. Integration tests demonstrate that although the components were individually satisfaction, as shown by successfully unit testing, the combination of components is correct and consistent. Integration testing is specifically aimed at exposing the problems that arise from the combination of components.

Functional test

Functional tests provide systematic demonstrations that functions tested are available as specified by the business and technical requirements, system documentation, and user manuals.

Functional testing is centered on the following items:

Valid Input : identified classes of valid input must be accepted.

Invalid Input : identified classes of invalid input must be rejected.

Functions : identified functions must be exercised.

Output : identified classes of application outputs must be exercised.

Systems/Procedures : interfacing systems or procedures must be invoked.

Organization and preparation of functional tests is focused on requirements, key functions, or special test cases. In addition, systematic coverage pertaining to identify Business process flows; data fields, predefined processes, and successive processes must be considered for testing. Before functional testing is complete, additional tests are identified and the effective value of current tests is determined.

System Test

System testing ensures that the entire integrated software system meets requirements. It tests a configuration to ensure known and predictable results. An example of system testing is the configuration oriented system integration test. System testing is based on process descriptions and flows, emphasizing pre-driven process links and integration points.

White Box Testing

White Box Testing is a testing in which in which the software tester has knowledge of the inner workings, structure and language of the software, or at least its purpose. It is purpose. It is used to test areas that cannot be reached from a black box level.

Black Box Testing

Black Box Testing is testing the software without any knowledge of the inner workings, structure or language of the module being tested. Black box tests, as most other kinds of tests, must be written from a definitive source document, such as specification or requirements document, such as specification or requirements document. It is a testing in which the software under test is treated, as a black box .you cannot "see" into it. The test provides inputs and responds to outputs without considering how the software works.

Unit Testing

Unit testing is usually conducted as part of a combined code and unit test phase of the software lifecycle, although it is not uncommon for coding and unit testing to be conducted as two distinct phases.

Test strategy and approach

Field testing will be performed manually and functional tests will be written in detail.

Test objectives

- All field entries must work properly.
- Pages must be activated from the identified link.

Test Case 1	
Test Case Name	Empty login fields testing
Description	In the login screen if the username and password fields are empty
Output	Login fails showing an alert box asking to enter username and password.

- The entry screen, messages and responses must not be delayed.

Features to be tested

- Verify that the entries are of the correct format
- No duplicate entries should be allowed

Test Case 3	
Test Case Name	User application success
Description	User login need to provide all data.
Output	Enter valid application

- All links should take the user to the correct page.

IntegrationTesting

Software integration testing is the incremental integration testing of two or more integrated software components on a single platform to produce failures caused by interface defects.

The task of the integration test is to check that components or software applications, e.g. components in a software system or – one step up – software applications at the company level – interact without error.

Test Results: All the test cases mentioned above passed successfully. No defects encountered.

Acceptance Testing

User Acceptance Testing is a critical phase of any project and requires significant participation by the end user. It also ensures that the system meets the functional requirements.

Test Results: All the test cases mentioned above passed successfully. No defects encountered.

CONCLUSION :

Voice-Controlled Devices uses Natural Language Processing to process the language spoken by the human to understand and process the query and respond to the human with the result. It eases most of the tasks of the user like searching the web and opening applications on desktop. In the process of making this project we realised that the concept of Artificial Intelligence in every field is decreasing human efforts and saving a lot of time.

FUTURE SCOPE:

• Make SRIDEVI to learn more on its own and develop a new skill in it.
• SRIDEVI android app can also be developed.
• SRIDEVI is also expected to take a more proactive role.
• Voice commands can be encrypted to maintain security.

REFERENCES:

1. Harshit Agrawal, Nivedita Singh, Gaurav Kumar, Dr. Diwakar Yagyasen, Mr. Surya Vikram Singh. "Voice Assistant Using Python" An International Open Access-revied, Refereed Journal.Unique Paper ID: 152099, Publication Volume & Issue: Volume 8, Issue 2, Page(s): 419-423.

2. George Terzopoulos, Maya Satratzemi "Voice Assistants and Smart Speakers in Everyday Life and In Education", Department of Applied Informatics, University of Macedonia, Thessaloniki, Greece.

3. Deepak Shende. Ria Umabiya, Monika Raghorte, Aishwarya Bhisikar. Anup Bhange. "Al Based Voice Assistant Using Python", International Journal of Emerging Technologies and Innovative Research (www.jetir.org), ISSN 2349-5162, Vol.6, Issue 2, page no.506-509, February-2019.

4. Tulshan, Amrita & Dhage, Sudhir. (2019). "Survey on Virtual Assistant: Google Assistant, Siri, Cortana, Alexa", 4th International Symposium SIRS 2018, Bangalore, India, September 19–22, 2018, Revised Selected Papers. 10.1007/978-981-13-5758-9_17.

5. Dr. Kshama V. Kulhalli, Dr.Kotrappa Sirbi, Mr. Abhijit J. Patankar, "Personal Assistant with Voice Recognition Intelligence", International Journal of Engineering Research and Technology. ISSN 0974-3154 Volume 10, Number 1 (2017).