**Desktop Voice Assistant**

Ramya Gujjala(Assistant Professor)  
*Electronics and Communication Engineering*  
*Bhojreddy Engineering College for Women,Hyderabad,India*

Soumya Koratpally  
*Electronics and Communication Engineering*   
*Bhojreddy Engineering College for Women,Hyderabad,India*

Nishitha Murki  
*Electronics and Communication Engineering*  
*Bhojreddy Engineering College for Women,Hyderabad,India*

Akhila Adepu  
*Electronics and Communication Engineering)*  
*Bhojreddy Engineering College for Women,Hyderabad,India*

Deepa  
*Electronics and Communication Engineering*   
*Bhojreddy Engineering College for Women,Hyderabad,India*

**ABSTRACT:**Owning a business, many of our important tasks can be neglected due to a busy schedule. In such circumstances, having a Voice Assistant will always come in handy as they can help you with many tasks. A remarkable increase in the development and deployment of voice-controlled Personal Digital Assistant has been seen in the recent years. The amazing feature of these assistants in common is their ability to interact with the human through oral communication. It takes the user input in form of voice and processes it and returns the output in various forms like action to be performed or the search result is dictated to the end user. This application model enables the user to control installed applications, send and receive e-mails or perform internet browsing. The model is also allows the users accessing their home appliances over voice commands. Speech recognition is a kind of technology that is using computer to transfer the voice signal to an associated text or command by identification and understand. This paper is all about the technology used in Voice Assistant i.e., Speech Recognition, its Synthesis, future aspects and scope.

**KEYWORDS*:***Voice Assistant, Speech Recognition***.***

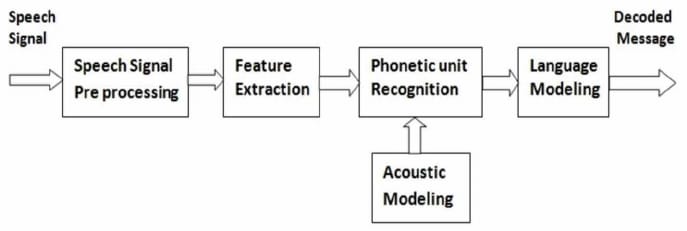
***1* .INTRODUCTION**:

Every system around us is becoming automatic and making our lives more comfortable. The underlying aim of every technology is to ease human efforts as far as possible, so with thisobjective in mind we set off to achieve another such goal. Voice Assistants are known for being great at many things and can make a positive contribution to the success of your small media enterprises(SME). They can also assist with your organization skills and productivity by freeing up your time for you to focus on other important activities.The rapid evolution of AI and machine learning made possible the development of voice recognition technology, which is how actively penetrating every area of our lives .A virtual desktop assistant which can perform different tasks without any physical intervention with keyboard and mouse. The system takes the voice commands of the user, analyzes the commands using speech recognition technology and acts accordingly.For the voice assistant all you need is concepts of python. A code which is used to understand our need is transformed accordingly.

**2. OVERVIEW OF ASR SYSTEMS:**

**ASR** stands for **Automatic Speech Recognition**.Speech Recognition is a kind of technology which is used by a computer to transfer the Speech signal to an associated text. As we have seen some special **Voice Assistant Like Amazon’s Alexa,Google Assistant,Microsoft’s Cortona and Apple’s Siri** .**Microsoft SAPI 5(Speech Application Programming Interface) is an API developed by Microsoft to allow the use of speech recognition and speech synthesis within Windows applications which is used to perform voice assistant.**It is not cloud-based ,so it can work also without Internet connection.It controls all the tasks.

Voice Assistant is a task-oriented programming application that recognizes human speech and carries out commands pronounced by a user .It is powered by Artificial Intelligence (AI).It is based on performance on cloud storage with millions of words and phases in it. Majorly, Voice Assistant Applications work based on Automatic Speech Recognition systems.This Voice Technology is used mostly in our daily life.This type of advancements increases day by day.



An Automatic Speech Recognition (ASR) has **2** important tasks:

1. **Phonetic Recognition**

2. **Word Decoding**.

ASR systems record the speech and then break it down into phonemes which are later get translated to text.

**2.1.Phonemes**:It is a basic unit of measurement for human speech recognition.

Firstly speech signal is sent,then it is processed.

**2.2.Feature Extraction:**Useful features are extracted from the speech signal on the basis of prior knowledge.This phase is Feature Extraction.

In a nutshell,the process starts with the device gathering audio with the microphone.Recorded speech waveforms get straight to **Acoustic Analysis**,which is performed on **3** different levels.

**2.2.1.Acoustic Modelling**:It represents which phonemes should be pronounced.

**2.2.2.Pronunciation Modelling**:This model will analyze the way phonemes are pronounced and also checks is there any accent or any other peculiarities of the vocal apparatus to capture the phonetic varialbility of speech.

**2.2.3.Language Modelling**:It is aimed at finding contextual probabilities depending on what phonemes were captured.

The speech waveforms data is then transmitted to the **decoder**,where it finally transforms into text for further use like command.

In this way Voice is converted to a text format.

**3.HOW SPEECH RECOGNITION WORKS?**

Let us know about Technology used in our Voice assistant i.e., Speech Recognition.

Speech Recognition means basically the Conversion of Speech to Text format .For instance, when you search something in Google using your voice, your voice is converted to text and later it is being searched .

When you speak something, sound waves are converted into digital ones and these digital ones are broken apart into very small pieces. These small pieces are then later compared with phonemes present in the database of a particular software. What the phonemes are? They are the smallest sounds present in a language .In English, there are nearly 40 phonemes and every word can be created using these 40 phonemes .So, when the small digital signals are compared with phonemes and when every small digital signal is assigned with phoneme, a sentence of phoneme is created .Then these phonemes are related with particular texts regarding them and then the whole text is created.

**4.LITERATURE SURVEY:**

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.The ASR and Search components perform speech recognition and search tasks.

In this literature review, the most information was available on Speech recognition technology, its applications, have insights of Bing Speech API and **Microsoft SAPI 5.**

The reviewed literature hasthe current state- of-the-art in assistive technology using Voice Recognition and gave the information required to develop a voice-controlled desktop assistant model on Microsoft Windows platform.

From these research papers we had known the importance of Speech Recognition.Speech recognition system will be more widely used for research in the future. A variety of speech recognition products would appear in the market.

Voice recognition would soon be a feature of every small to big devices ranging from ticket selling devices to washing machines e.tc. Set of techniques for improving the performance of automated voice search services intended for mobile users accessing these services over a range of portable devices.

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 potentially very large application specific database.

Study provides a good example of how additional domain specific knowledge sources can be used with a domain independent ASR system to facilitate voice access to online search indices. A detailed explanation of technology used that is, Speech Recognition is mentioned in the research papers.

#### 5. CIRCUIT DIAGRAM:

#### Screenshot (23).png

#### Hardware components:

* [Bolt WiFi Module](https://www.boltiot.com/cs/c/?cta_guid=46c823fb-bead-4c43-baf4-4af76be82181&placement_guid=f442d65c-fe31-43a3-b178-6a9f749125dd&portal_id=4801849&canon=https%3A%2F%2Fwww.boltiot.com%2Fprojects%2Fhome-automation&redirect_url=APefjpEho-D70knG7peXFT89dGbrgUerzslxPsvDnRyR5EVMadesebro0W8HKpKGVWRZxqfX6p3DOXFIyeK05ljW8-aBWs-XbNB4Qe20k0_kpoPHGCRXwDg&click=8b942e02-4c6b-42b5-88d9-48a50a386591&hsutk=6abf5dca25ab9e2a02503f6417a7f6ca&signature=AAH58kFZu5-Z2t6hZ8NyaCD7w-Y8j6fg8Q&pageId=7147045755&__hstc=155166375.6abf5dca25ab9e2a02503f6417a7f6ca.1600435358951.1600435358951.1600435358951.1&__hssc=155166375.1.1600435358952&__hsfp=4253523528&contentType=blog-post)
* Relay Module (Generic)
* Jumper wires (generic)
* Bulb
* Bulb holder
* Mobile charger

### Hardware Setup:

### Step 1: Gathering all the required components

1. Bolt Wifi Module

2. Single Channel Relay Module

3. Jumper Wires and normal wires

4. Bulb holder

5. Bulb

6. Mobile Charger

7. 2-Pin Socket

#### Step 2: Assembling the hardware

1. Connect the 5V supply pin and GND of the Bolt to the 5V and GND pins of the relay module.

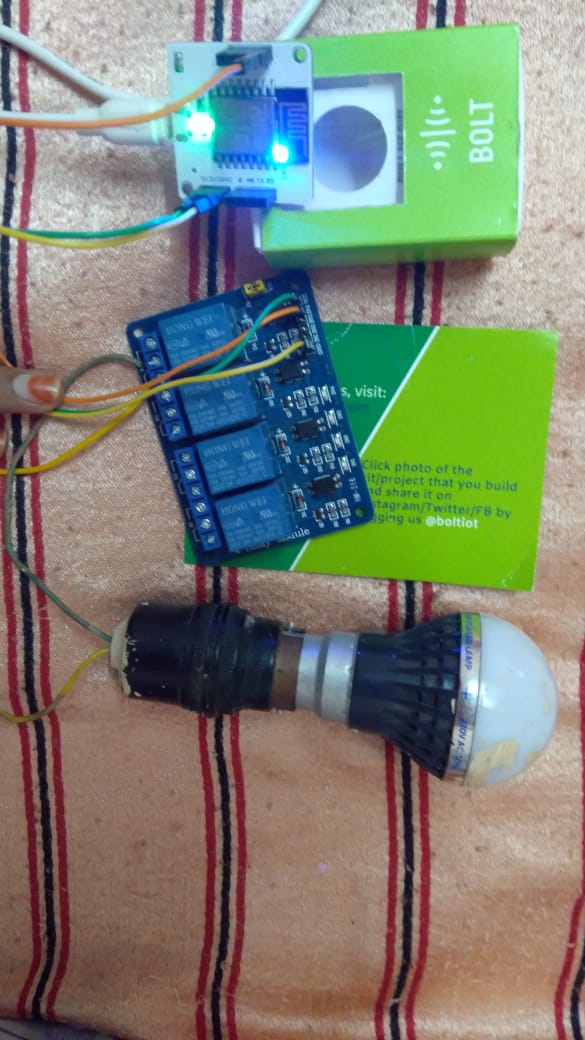
2. Connect any one of the digital pin i.e (0-4) of the Bolt to the IN/REL pin of the relay module.

3. Connect a wire from the bulb holder to the COM pin of the relay module.

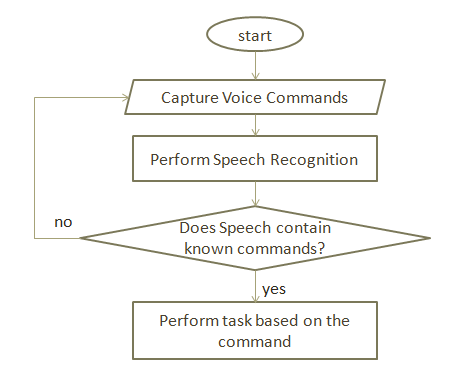
4. Connect a wire from the NO pin of the relay module to the 2-pin socket.

5. Connect the remaining terminal of the bulb holder tot the other terminal of the 2-pin socket

**Step 3: Connecting bolt to cloud**

******

**6. FLOW CHART:**

****

1).After initializing the assistant, it starts capturing the voice commands from the user .The custom voice commands are added to the code for specific command execution.

2).After capturing the voice commands the speech recognition was done .Automatic Speech Recognition system breaks the captured voice command into phonemes and then to Acoustic modelling.

3).After recognition process, if the user voice input has known commands, the respective tasks were performed. If the input doesn't have any known commands or else if the system cannot recognize the voice command it again starts capturing the input.

**7. FUTURE SCOPE:**

Voice Assistant based Artificial Intelligence has changed the lifestyle of people .We are talking about the next generation of tools to spark growth in retail, logistics, healthcare, smart cities, manufacturing, and autonomous vehicles, among many others .Voice technology is increasing day by day.This is only Google has made moves in making Assistant more ubiquitous by opening the software development kit through [Actions](https://developers.google.com/actions/), which allows developers to build voice into their own products that support artificial intelligence. Another one of Google’s speech-recognition products is the AI-driven [Cloud Speech-to-Text](https://cloud.google.com/speech-to-text/) tool which enables developers to convert audio to text through [deep learning](https://www.techemergence.com/deep-learning-in-the-enterprise-current-traction-and-challenges/)neural network algorithms. In this way we can see many advancements can be made in voice technology in the future.

**8.RESULT:**

Our Voice assistant has the following features:

1)It can send messages through 'Twilio' to our phones.

2) It mails through 'Mailgun'.

3) It has Wikipedia search.

4) It writes a note of what we say and stores it in a Seperate Folder.

5) It plays Youtube.

6)We can access the file through its path given.

7)It has an home automation system which turns on and off the lights.

8) It says time.

9) If the user not accessing any of the above then it directly get response from the chatbot present within the commands.

10) Plays music.

11) Greets you.

12)Shutdown.

13)Restart and many more.

**9.CONCLUSION:**

Speech Recognition is a wonderful application of AI.It has major applications on speech.But this Application mostly works with Internet Connections. But our Proposed System has capability to work with and without Internet Connectivity. It is named as 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. In addition, this proposed system can change the way of interactions between end user and the PC’s. Through this paper an attempt has been made to provide a review of how this technology actually works and how far the progress had been made in past few years. Computers would quickly arrive with preinstalled automatic speech recognition systems.Desktop Assistant has various functionalities, and works only based on the voice commands. This is language barrier independent,we can develop the system in various languages. It actively responds to user’s voice commands.

The Assistant developed using python and Speech Recognition Technology. Speech Recognition is an excellent application of artificial intelligence to work with since speech is one of the most basic activities of humans.The system uses a simple use of API (Application Program Interface) call that turns the state of the digital pin to HIGH or LOW. The state of this pin is further used to switch the relay ON or OFF. And this API call is activated when you press the button on the browser. When the relay is in OFF state it breaks the circuit between the Bulb and the power supply result of which the bulb goes off whereas when the relay is in ON state it completes the circuit connection between bulb and power supply as a result of which bulb lights up.

**10. REFERENCES:**

1. Digital Life Assistant using Automated Speech Recognition,2014. International Conference on Innovative Applications of Computational Intelligence on Power, Energy and Controls with their Impact on Humanity (CIPECH14).
2. [https://](https://www.ripublication.com/irph/ijert_spl17/ijertv10n1spl_80.pdf)[www.ripublication.com/irph/ijert\_spl17/ijertv10n1spl\_80.pdf](http://www.ripublication.com/irph/ijert_spl17/ijertv10n1spl_80.pdf). International Journal of Engineering Research and Technology. ISSN 0974-3154 Volume 10, Number 1 (2017) © International Research Publication House
3. Personal Assistant with Voice Recognition Intelligence,2017. International Journal of Engineering Research and Technology. ISSN 0974-3154 Volume 10, Number 1 (2017) © International Research PublicationHouse<http://www.irphouse.com>
4. Marcus E.Hennecke , K.Venaktesh Prasad, David G. Stork , “Automatic Speech Recognition System using Acoustic & Visual Signals” , 1996 IEEE Proceedings of ASILOMAR-29.
5. ShrutikaKhobragade, “Jarvis Digital Life Assistant”, international journal of engineering research and technology-volume 2 issue 1-2013.
6. Preeti Saini , Parneet Kaur, “Automatic Speech Recognition” , international journal of engineering tools and technology-volume 4 issue 2-2013.
7. <http://www.ieeexplore.ieee.org>
8. Marcus E.Hennecke , K.Venaktesh Prasad, David G. Stork , “Automatic Speech Recognition System using Acoustic & Visual Signals” , 1996 IEEE Proceedings of ASILOMAR-29.
9. RaghvendraPriyam, Rashmi Thakur, VidehKishori Thakur, “Artificial Intelligence Applications for Speech Recognition”, CAC2S 2013.
10. ShrutikaKhobragade, “Jarvis Digital LifeAssistant”, international journal of engineering research and technology-volume 2 issue 1-2013.
11. Hector Perez-Meana, Advances in audio and speech signal processing: technologies and applications, US: Idea Group Publishing, 2007.
12. Rabiner L R, Juang B H. Fundamentals of speech recognition, Englewood Cliffs: Prentice Hall, 1993

**13.** “Language modeling for What-with Where on GOOG411,” Proc. of INTERSPEECH, 2009