Mini-Project Report

on

**Voice based Text-input system**

(CSE IV Sem Mini Project)

 (2021-2022)

Submitted By:

Swayam Manori

Student ID: 20011825

CSE-C-IV Sem

Session: 2021-2022

Submitted To:

**Mr Dilip Gangwar**

Associate Professor

GEHU, D.Dun

   



Certifies that Swayam Manori (Student ID.- 20011825) has

developed a mini project on **“Voice Based Text Input System”** for

the CS IV Semester Mini Project in Graphic Era Hill University,

Dehradun. The project was carried out by Student is their own work

as best of my knowledge.

Date: 12th July,2022

**Mr Dilip Gangwar**

CSE-B-IV-SEM

(CSE Department)

GEHU, Dehradun



We would like to express our gratitude to The Almighty Shiva Baba,

the most Beneficent and the most Merciful, for completion of project.

We wish to thank our parents for their continuing support and encouragement. We also wish to thank them for providing us with the opportunity to reach this far in our studies.

We would like to thank our class coordinator for his patience, support and encouragement throughout the completion of this project and having faith in us.

We also acknowledge them who help us in developing the project.

At last, but not the least, We are greatly indebted to all other persons who directly or indirectly has helped us during this work.

**Mr. Swayam Manori**

**Student ID.- 20011825**

**CSE-C-IV-SEM**

**Session: 2021-2022**

**Table Of Contents**

1. **Introduction** 
   1. Introduction
   2. Amazon’s Echo Dot
   3. About the Project.
2. **Requirements For Projects**
   1. Software Requirements
   2. Permissions to be Granted
   3. Libraries Installed
3. **Designing The FrontEnd** 
   1. HTML Part
   2. CSS Part
4. **Designing the Back-End** 
   1. JavaScript Speech Recognition Module
   2. Event Handlers
5. Conclusion
6. References



Speech To Text Recognition is a recognition Software that provides the facility to recognize and translate the spoken language into text using certain computer linguistics. It is also known as Speech Recognition. Speech Recognition has quickly transcended from everyday use on mobile phones to our everyday lives like in banking, medical etc. Speech recognition applications reveal how voice to text technology can increase the efficiency of simple tasks and extend to tasks that humans have traditionally performed. The massive eye-catching moment for Speech recognition came in 2008 via J.A.R.V.I.S., Tony Stark’s Virtual Butler, in the science fiction movie Iron Man. J.A.R.V.I.S. started as a computer interface which was then upgraded to an A.I. program which ran the business.

J.A.R.V.I.S. opened the eyes and ears of the world to the possibilities to innate in Speech Recognition technology in various fields. Speech recognition provides us with hands free control on our devices like speakers, smartphones and even vehicles in a wide variety of languages.

# Amazon’s Echo Dot:

Icon

Description automatically generated

One of the most powerful Speech Recognition devices, which is used world-wide in almost everyone’s home, Amazon’s Echo Dot. It is a voice-controlled Speaker that uses amazon’s artificial Intelligent personal assistant, Alexa.

Echo Dot can play music, answer any of your questions, set the volume to your desired value, set alarms, and even connect to other Smart devices in your home, via the use of Alexa.

**About The Project:**

As the name suggests, voice-based text input system lets a user input text using his speech. Or in laymen language we can say that it is a speech to text convertor which listens to a voice in a certain language(in this case English) and writes the spoken text in a text area.

This comes under a machine learning project as we are using a in-build module of JavaScript which is The Speech recognition Module. The project is made with the help of HTML, CSS and JavaScript.

The Project runs on a website, where there is a button which upon click, records everything the user say via the system’s or an external microphone and displays the text in the text area. In addition to this, it also performs functions such as Searching of a desired keyword, opening google docs, slides, sheets, and forms and searching a location. All these functionalities using voice commands only.

The webpage is developed and styled using the Markup Language And Styling Language- HTML and CSS respectively. The Speech Recognition is applied using the scripting language-JavaScript.



**Software Requirements:**

1. A latest version of Visual Studio Code.
2. Latest version of a Web Browser (Mozilla, Chrome, MS Edge).

**Permission to be Granted:**

1. Microphone Permission (only when running for the first time).

**Libraries Installed:**

1. A Live Server to launch your .html and make live changes without saving.
2. Bootstrap loader for CSS or directly copy paste the link of CSS from bootstrap website.
3. Speech Recognition Module in JavaScript.

**Speech Recognition:**

As mentioned earlier, speech recognition is the capability to process human speech into written format. Speech recognition works by listening to the vibrations from our voice. The vibrations are then converted to digital form via an analog to digital convertor. The sounds are then segmented into hundredths or thousandths of seconds and are then matched to phonemes. A phoneme is a unit of sound that distinguishes one word from another in any given language. For example, there are approximately 40 phonemes in the English language. The phonemes are then run through a network via a mathematical model that compares them to well-known sentences, words, and phrases. The text is then presented as text.

A picture containing text, clipart

Description automatically generated

Diagram

Description automatically generated

**Designing the Front-End**

The system consists of a simple webpage containing a heading, 2 buttons and a text area where the text will be inputted as per the voice.

* HTML and CSS is used to create all 4 components.
* The text area and buttons are given a id so that we can access them in JavaScript.
* The Speech.js contains our backend portion which is the Speech Recognition Module.

Text

Description automatically generated

**Designing the Back-End**

The backend is made using the speech recognition module of JavaScript.

The Module contains many event handlers which will help us to record how different actions will be performed like what to do when user stops speaking or if there’s an error.

* First of all, we select the textbox and button via their respective ids.
* Now we create a object of SpeechRecognition.
* Now to add all the event handlers one by one.

Text

Description automatically generated

* The onstart event handler will work when the start button will be clicked and so we

add a action event on button which will be:

Graphical user interface, text

Description automatically generated

* Now, we’ll we what will happen when the recognition will be stopped.

Text

Description automatically generated

Graphical user interface, text

Description automatically generated with medium confidence

Text

Description automatically generated

* The three Event Handlers are provided to some particular tasks when the user has stopped speaking.

1. Onspeechend will stop the recognition and clear the textbox.
2. Onaudioend will display a message on the console that audio has ended.
3. Finally onend will check that if speechrecognitionison variable is true,ifyes, it will continue to recognize else it will end the recognition.

* Now, we will look how to input text in the textbox. This is achieved via onresult event handler.
  + We first extract the result from SpeechRecognitioResult and SpeechRecognitionAlternative objects and store it in the variable. We then store the result in the content variable and place the content’s value in the textbox.
  + The code is as follows ;

Text

Description automatically generated

**Conclusion**

We have successfully developed our Voice Based Text Input System project. In this, we have created a webpage using HTML CSS and used a Speech Recognition Module of JavaScript with recognition accuracy of 80%.

Voice bases Text Input System is very useful in today’s era where the world is moving towards the metaverse. It has various benefits such as:

* Save time: Automatic speech recognition technology saves time by delivering accurate transcripts in real-time.
* Cost-efficient: Most speech to text software has a subscription fee, and a few services are free. However, the cost of the subscription is far more cost-efficient than hiring human transcription services.
* Enhance audio and video content: Speech to text capabilities mean that audio and video data can be converted in real-time for subtitling and fast video transcription.
* Streamline the customer experience: By drawing on natural language processing, the customer experience is transformed through ease, accessibility, and seamlessness.

**References**

* 1. <https://aws.amazon.com/what-is/speech-to-text/#:~:text=Speech%20to%20text%20is%20a,recognition%20or%20computer%20speech%20recognition.>
  2. <https://www.youtube.com/watch?v=rwB6RqqCmXc>.
  3. <https://overapi.com/css>
  4. <https://overapi.com/javascript>