REAL-TIME COMMUNICATION SYSTEM POWERED BY AI FOR SPECIALLY ABLED

A PROJECT REPORT

Submitted by

Team ID: PNT2022TMID21808

RAJESH. R	142219106074
LOGAVARTHAN. T	142219106048
MYTHRAYAN. S	142219106060
RAJASEKAR. G	142219106073

TABLE OF CONTENTS

1.	**INTRODUCTION**	**1**
	1.1. Project Overview	**1**
	1.2. Purpose	**1**
2.	**LITERATURE SURVEY**	**2**
	2.1. Existing Problem	**2**
	2.2. References	**4**
	2.3. Problem Statement Definition	**5**
3.	**IDEATION & PROPOSED SOLUTION**	**6**
	3.1. Empathy Map Canvas	**6**
	3.2. Ideation & Brainstorming	**7**
	3.3. Proposed Solution	**10**
	3.4. Problem Solution fit	**13**
4.	**REQUIREMENT ANALYSIS**	**15**
	4.1. Functional requirement	**15**
	4.2. Non-Functional requirements	**16**
5.	**PROJECT DESIGN**	**17**
	5.1. Data Flow Diagrams	**17**
	5.2. Solution & Technical Architecture	**19**
	5.3. User Stories	**23**
6.	**PROJECT PLANNING & SCHEDULING**	**24**
	6.1. Sprint Planning & Estimation	**24**
	6.2. Sprint Delivery Schedule	**24**

6.3. Reports from JIRA **26**

7. CODING & SOLUTIONING **28**

7.1. Feature 1 **28**

7.2. Feature 2 **33**

8. TESTING **35**

8.1. Test Cases **35**

8.2. User Acceptance Testing **35**

9. RESULTS **37**

9.1 Performance Metrices **37**

10. ADVANTAGES & DISADVANTAGES **40**

11. CONLUSION **41**

12. FUTURE SCOPE **41**

13. APPENDIX **42**

13.1 Source Code **42**

13.2 GitHub & Project Demo Link **44**

CHAPTER 1

INTRODUCTION

1.1 Project Overview

Technology is developing day by day but no significant developments are performed for the improvement of the people with disabilities in communication like deaf and dump. Communication between the deaf and a normal person has always been a challenging task. It is very difficult for dumb people to get their message across to normal people. Since ordinary people are not trained in sign language. Sign language is the only tool of communication for the person who is not able to speak and hear anything. Sign language is a boon for the physically challenged people to express their thoughts and emotion. In this work, a novel scheme of sign language recognition has been proposed for identifying the alphabets and gestures in sign language. We are making use of a convolution neural network to create a model that is trained on different hand gestures. An app isbuilt which uses this model. This app enables deaf and dumb people to convey their information using signs which get converted to human-understandable language and speech is given as output. It bridges the communication gap between people who cannot speak and the general public

1.2 Purpose

The 2011 Indian census cites roughly 1.3 million people with “hearing impairment”. In contrast to that numbers from India’s National Association of the Deaf estimates that 18 million people – roughly 1 per cent of Indian population are deaf. These statistics formed the motivation for our project.

- As these speech impairment and deaf people need a proper channel to communicate with normal people there is a need for a system.
- Not all normal people can understand sign language of impaired people.
- Normal people face difficulty in understanding their language. Hence there is a need of a system which recognizes the different signs, gestures and conveys the information to the normal people.

CHAPTER 2

LITERATURE SURVEY

In this chapter, the similar previous projects that developed by others is reviewed and discussed to give critical review and also to choose the suitable equipment and method before starting the project implementation. The chapter summaries of each of the project works are mentioned below

2.1 Existing Problem

The verbal exchange among a dumb and listening to individual poses to be an essential drawback in comparison to verbal exchange among blind and historic visible human beings. This creates an exceedingly little residence for them with verbal exchange being accomplice diploma basic issue of human life. The blind human beings can communicate freely through means that of historic language while the dumb have their very own manual-visible language called signal language. Sign language is likewise a non-verbal symbolization that is observed amongst deaf groups at periods the planet. The signal languages have not been given a standard foundation and as a result tough to interpret. A Dumb verbal exchange interpreter is likewise a device that translates the hand gestures to sensibility speech. A gesture in accomplice diploma extraordinarily language is likewise a positive motion of the arms with a specific type created out of them. A gesture in a signal language is a specific motion of the arms with a particular form produced from them. A signal language typically presents signal for entire phrases. It also can offer signal for letters to carry out phrases that do not have corresponding register that signal language. In this tool Flex Sensor performs the fundamental role, Flex sensors are sensors that alternate in resistance relying on the quantity of bend at the sensor.[1]

This virtual glove targets to decrease this barrier in communique. It is digital tool which can translate Sign language into speech to be able to make the communique take area among the mute groups with the overall public possible.[2]

A hand gesture recognition system is likewise used to apprehend actual time gesture in unconstrained environments. The gadget includes 3 modules: actual time hand monitoring, education gesture and gesture popularity the usage of pseudo size hidden Markov models. In this they've used a Kalman clear out and hand blobs evaluation for hand monitoring to attain movement descriptors and hand region.[3].

These days advanced intensity sensors, e.g., the Kinect sensor, have furnished new possibilities for human computer interaction (HCI). Although high-quality development has been made through leveraging the Kinect sensor, e.g., in human frame tracking, face popularity and human motion popularity, strong hand gesture popularity stays an open trouble. Compared to the complete human frame, the hand is a smaller item with extra complicated articulations and extra without difficulty laid low with segmentation errors. It is as a result a totally hard trouble to understand hand gestures. This paper specializes in constructing a strong partprimarily based totally hand gesture popularity device the use of Kinect sensor.[4]

Sign language is a beneficial device to ease the conversation among the deaf man or woman and everyday man or woman. The machine ambitions to decrease the conversation hole among deaf humans and everyday world, because it helps way communications. The projected method interprets language into speech. The machine overcomes the necessary time problems of speech-listening to impaired and improves their manner. This machine converts the language in associate passing voice it's nicely explicable via way of means of deaf humans. With this undertaking the deaf-mute humans can use the gloves to perform signal language and it'll be transformed into speech.[5]

Chat applications have become a powerful media that assist people to communicate in different languages with each other. There are lots of chat applications that are used different people in different languages but there is not such a chat application that has facilitate to communicate with sign languages. The developed system is based on Sinhala Sign language. The system has included four main components as text messages are converted to sign messages, voice messages are converted to sign messages, sign messages are converted to text messages and sign messages are converted to voice messages. Google voice recognition API has used to develop speech character recognition for voice messages. The system has been trained for the speech and text patterns by using some text parameters and signs of Sinhala Sign language is displayed by emoji. Those emoji and signs that are included in this system will bring the normal people closer to the disabled people. This is a 2-way communication system but it uses pattern of gesture recognition which is not very reliable in getting appropriate output.[6]

Speech impairment is a disability which affects one’s ability to speak and hear. Such individuals use sign language to communicate with other people. Although it is an effective form of communication, there remains a challenge for people who do not understand sign language to

communicate with speech impaired people. The aim of this paper is to develop an application which will translate sign language to English in the form of text and audio, thus aiding communication with sign language. The application acquires image data using the webcam of the computer, then it is preprocessed using a combinational algorithm and recognition is done using template matching. The translation in the form of text is then converted to audio. The database used for this system includes 6000 images of English alphabets. We used 4800 images for training and 1200 images for testing. The system produces 88% accuracy. [7]

2.2 References

[1]. V.Padmanabhan, M.Sornalatha," Hand gesture recognition and voice conversion system for dumb people," International Journal of Scientific & Engineering Research, Volume 5, Issue 5, May-2014.

[2]. Praveenkumar S Havalagi, Shruthi Urf Nivedita M.Tech (VLSI), Department of ECE, SIET, Bijapur, India B.E (ECE), Department of ECE, SIET, Bijapur, India," The Amazing Digital Gloves that give voice to the voiceless," International Journal of Advances in Engineering & Technology, Mar. 2013

[3]. Mitra, Senior Member, IEEE, and Tinku Acharya, Senior Member, IEEE," Gesture Recognition: A Survey" IEEE TRANSACTIONS ON SYSTEMS, MAN, AND CYBERNETICS—PART C: APPLICATIONS AND REVIEWS, VOL. 37, NO. 3, MAY 2007.

[4] Zhou Ren, et.al, IEEE" Robust Part Based Hand Gesture Recognition Using Kinect Sensor," IEEE TRANSACTIONS ON MULTIMEDIA, VOL. 15, NO. 5, AUGUST 2013.

[5]. Shraddha R. Ghorpade , Prof. Surendra K. Waghmare "A Communication System for Deaf and Dumb People",International Journal of Science and Research (IJSR) ISSN (Online): 2319-7064 Index Copernicus Value (2013).

[6]. P. G. Ahire, K. B. Tilekar, T. A. Jawake and P. B. Warale, "Two Way Communicator between Deaf and Dumb People and Normal People," 2015 International Conference on Computing Communication Control and Automation, 2015, pp. 641-644.

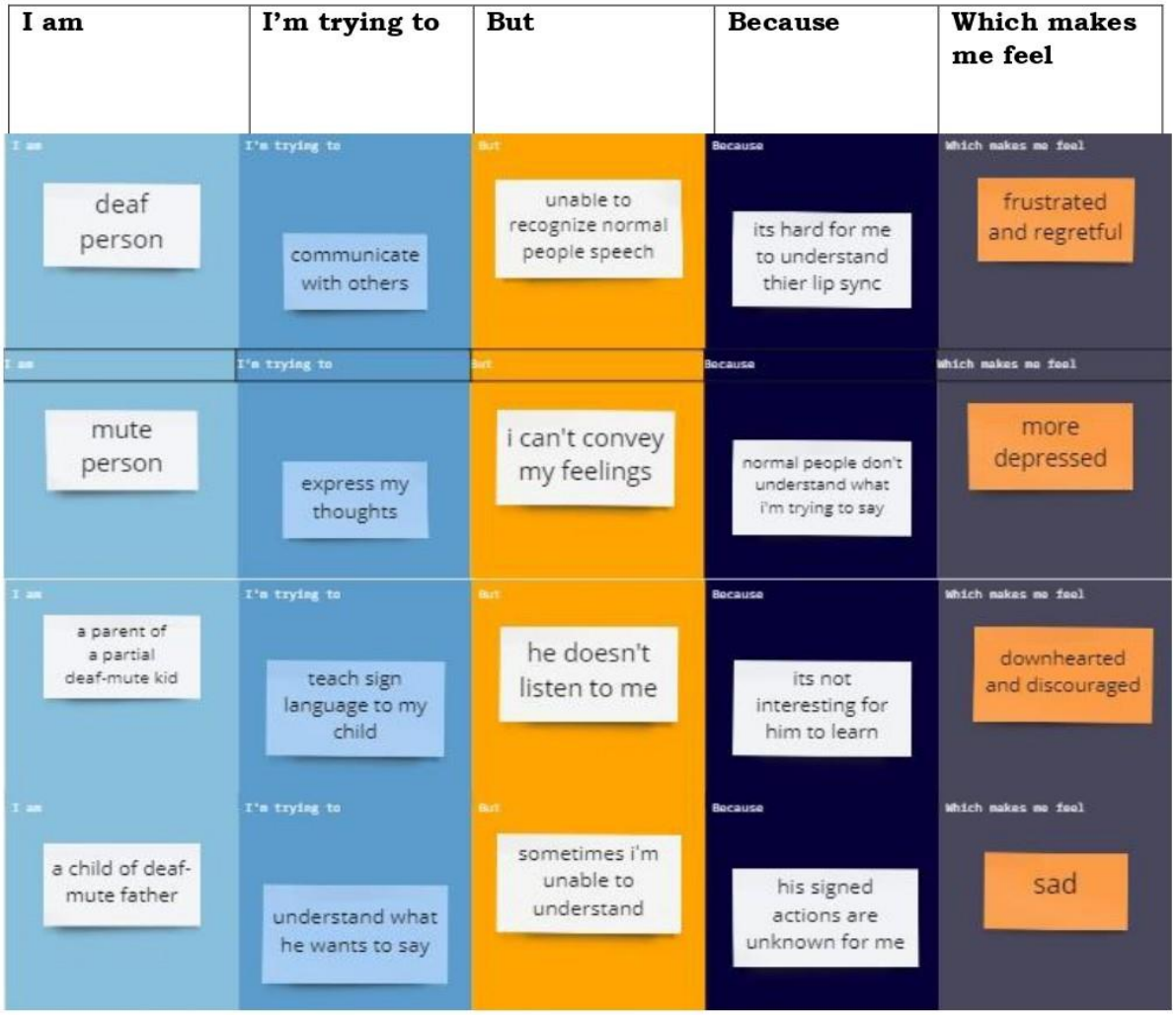
[7]. Gunasagari G. S, Poornima N, Abhijna Yaji, Achuth M, Anisha Maria Dsilva, Chethana S R, "Review on Text and Speech Conversion Techniques based on Hand Gesture", 2021 5th

International Conference on Intelligent Computing and Control Systems (ICICCS), pp.1682-1689, 2021.

2.3 Problem Statement Definition

Crafting a problem description helped us understand the customer's perspective. A well-formulated customer problem description helps to find the ideal solution to the challenges customers face. Throughout the process, you can also empathize with your customers, which will help you better understand how they perceive your product or service.

Here the table describes the customer and their characteristic, what they want to achieve, the problems that stand in their way, the reason for the problems, the emotion that results from the problem experienced.



CHAPTER 3

IDEATION & PROPOSED SOLUTION

The project aims to develop a system that converts the sign language into a human hearing voice in the desired language to convey a message to normal people, as well as convert speech into understandable sign language for the deaf and dumb. We are making use of a convolution neural network to create a model that is trained on different hand gestures. An app is built which uses this model. This app enables deaf and dumb people to convey their information using signs which get converted to human-understandable language and speech is given as output.

3.1 Empathy Map Canvas

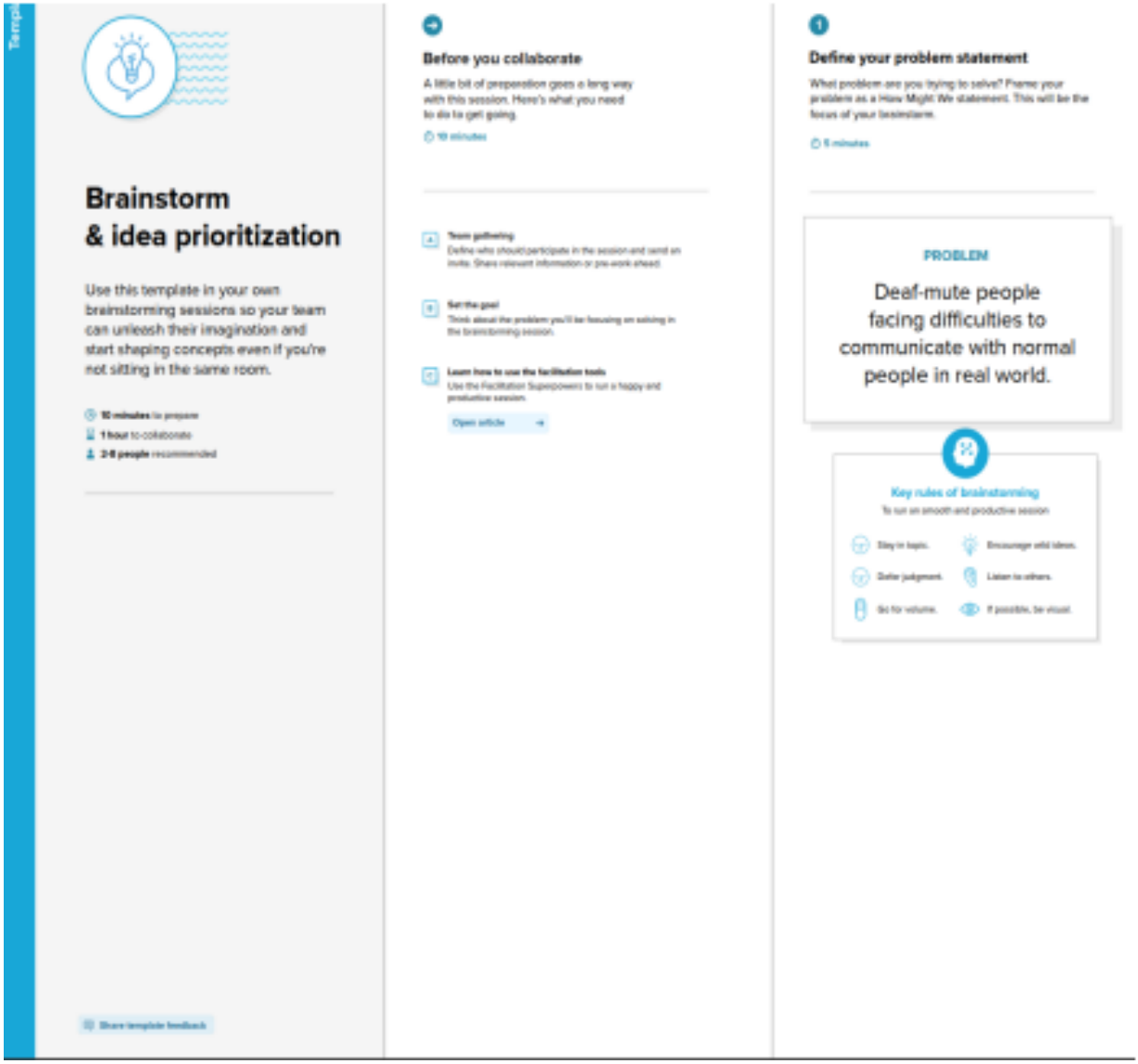
An empathy map is a simple, easy-to-digest visual that captures knowledge about user behaviors and attitudes. It is a useful tool to helps teams better understand their users. Creating an effective solution requires understanding the true problem and the person whois experiencing it.

3.2 Ideation & Brainstorming

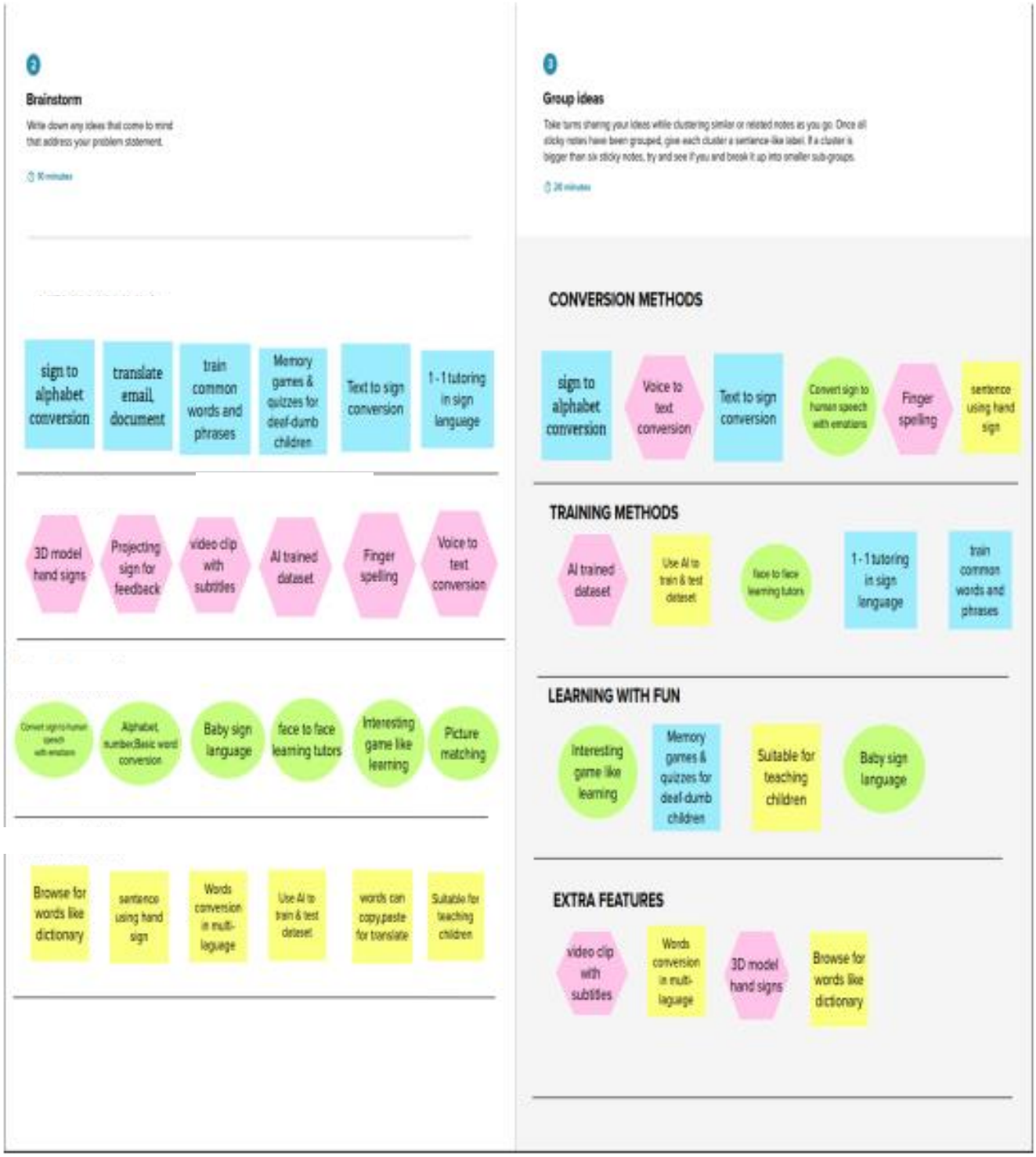
Brainstorming provides a free and open environment that encourages everyone within a team to participate in the creative thinking process that leads to problem solving. Prioritizing volume over value, out-of-the-box ideas are welcome and built upon, and all participants are encouraged to collaborate, helping each other develop a rich amount of creative solutions. Here in step 1, we gather our team and select the problem statement. Defining our problem statement help us to proceed for the next step.

Reference: https://www.mural.co/templates/empathy-map-canvas

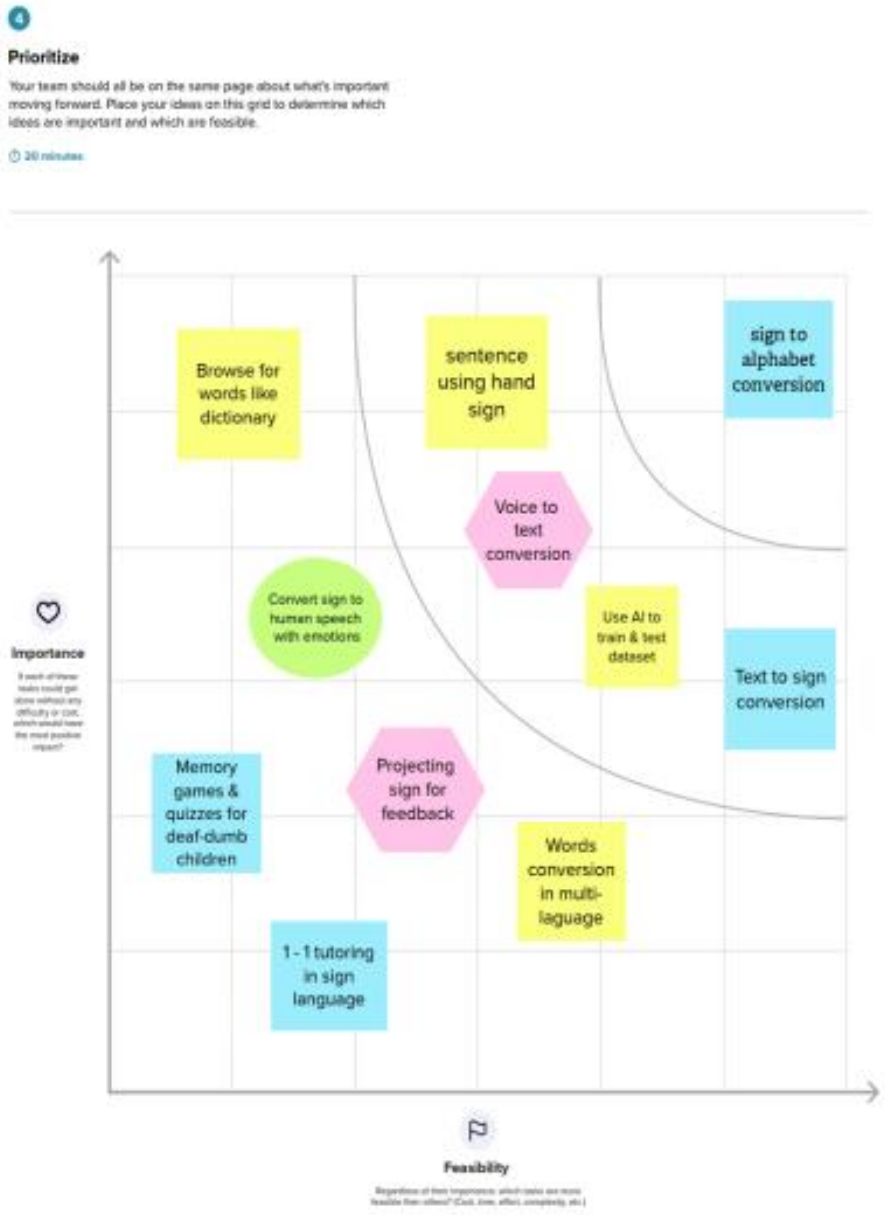
Step-1: Team Gathering, Collaboration and select the Problem Statement



Step-2: Brainstorm, Idea Listing and Grouping



Step-3: Idea Prioritization



3.3 Proposed Solution

Proposed Solution Template:

Project team shall fill the following information in proposed solution template.

S.No.	Parameter	Description
1.	Problem Statement (Problem to be solved)	➢ The deaf and dumb people have difficulty in communicating with others who cannot understand sign language and misinterpreted. ➢ They unable to communicate one's own thoughts, feelings, needs, and desires is incredibly frustrating and limiting. Parents and caregivers can also become frustrated when they areunable to understand their loved one. ➢ Uneducated Deaf-Dumb people face serious problem in Communication with normal people.
2.	Idea / Solution description	➢ An efficient and fast algorithm for identification of the number of fingers opened in a gesture representing an alphabet in the sign language. ➢ Sign Language into text using timefor deaf and dumb people.

		➢ Sign language application provide common signs, including those for letters, numbers, and everyday words. ➢ People can use to browse for words and phrases, some are interactive and offer games andquizzes to help people associate signs with their meanings more quickly.
3.	Novelty / Uniqueness	➢ Can be used for uneducated deaf-mute people. ➢ An interactive model for disable kids.
4.	Social Impact / Customer Satisfaction	➢ An integrated android application to blend uneducated Deaf-Dumb people within society, and help them tocommunicate with normal people. ➢ Deaf-Dumb people also havenormal opportunities for learning. ➢ Offer a great tool for parents to teach their deaf and dumb kids
5.	Business Model (Revenue Model)	➢ AI-based and image recognition systems in clouds, providing customers with Software as a Service and enabling your business to rapidly evolve and grow. ➢ Packages range from casual learning (6 hours) to unstoppablelearning (20 hours).

		➢ connect with certified teacherswho are Deaf/hard of hearing and allows you to learn and practice sign 1-1.
6.	Scalability of the Solution	➢ Our model has the capacity to handle growth, especially in handling more users and evolving concurrently with business needs (app's backend, database and the hosted servers). ➢ The solution is fully cloud- hosted, yielding scalability and on-demand service, without any requirement of the internal infrastructure or maintenance cost. ➢ The solution is easy tointegrate into another ecosystem due to the app’s modular structure.

3.4 Problem Solution Fit

1.CUSTOMER SEGMENT **Who is your customer?** ✓ Deaf people who eager to learn. ✓ Dumb (Mute) people who wants to expression their thought. ✓ Parents of deaf-mute kids. ✓ Children of disabled parents.	**6.CUSTOMER CONDTRAINTS** **What constrains prevent your customers from taking action or limit their choices of solution?** ✓ Network issues may delay to access the feature ✓ Improper sign language will lead the people to misunderstood.	**5.AVAILABLE SOLUTIONS** **Which solution are available to the customer when they face the problem or need to get the job done? What pro & cons do these solution have?** ✓ Sign to alphabet conversation ✓ Word conversation in multilanguage. ✓ Train common word and phrases Pros: ✓ Convert the sign into voice with emotion Cons: ✓ During sign to text conversation there may be an error. ✓ There may be chance of misunderstanding word in voice to text conversation.
2. JOBS-TO-BE DONE/PROBLEMS **Which jobs-to-be-done do you address for your customers?** ✓ No need of the translator for communication between deaf-mute people and normal people. ✓ Reduce the time to express their thoughts. ✓ Reduce cost used for translator.	**9. PROBLEM ROOT CAUSE** **What is the real reason that this problem exists? What is back story behind the need to do this job?** ✓ Deaf-mute is a birth disorder. ✓ Sometimes it occurs due to aging factor and accidents. ✓ Normal people are not so patient to understand deaf-mute people thoughts.	**7. BEHAVIOUR** **What does your customer do to address the problem and get the job done?** ✓ Taking the hand gesture image with goodquality and uploading it for conversion. ✓ Make use of text to voice conversion option while communicating with others.

3.TRIGGERS

What triggers customers to act?

- ✓ Customer doesn't need human translator to communicate with normal people.
- ✓ Instant result for the conversion of sign language to human understandable language.
- ✓ Interactive sessions.

10. YOUR SOLUTION

If you are working on an existing business
write down your current solution first fill in the canvas and check how much it fits reality

- ✓ Accurate conversion will be available
- ✓ Interactive and user-friendly solution to make it accessible
- ✓ Elimination of human error and fast functionality
- ✓ Provide common signs, including those for letters, numbers, and everyday words and phrases.

8. CHANNEL OF BEHAVIOR

8.1 ONLINE

What kind of actions do customers take online?

- ✓ Accessing required conversions using application.
- ✓ Uploading the image of hand sign.
- ✓ Quick access of Al based algorithm.

4. EMOTIONS: BEFORE/AFTER

How do customers feel when they face a problem or a job before and afterwards?

BEFORE:

- ✓ Feeling **MORE DEPRESSED** when unable express thoughts and feelings.
- ✓ Feeling **FRUSTRATED** and **REGRETFUL** while unable to recognize people speech
- ✓ Parents feeling **DOWNHEARTED** when can't teach their disable kids.

AFTER:

- ✓ Customers can feel **INDEPENDENT** and **CONFIDENT.**
- ✓ **UNPRESSURIZED** as
- ✓ they can share their thoughts.
- ✓ Parents,relatives and friends of disable people are **HAPPY**

8.2 OFFLINE

What kind of actions do customers take offline?

- ✓ Taking the hand sign picture properly without any blur.
- ✓ Parents make kids to practice in this application.
- ✓ Children of elderly disabled people helps to teach their parents.

CHAPTER 4

REQUIRMENTS ANALYSIS

Requirements analysis, also called requirements engineering, is the process of determining user expectations for a new or modified product. These features, called requirements, must be quantifiable, relevant, and detailed. In software engineering, such requirements are often called functional specifications. Requirements analysis is critical to the success or failure of a systems or software project..

4.1 Functional Requirements

FR No.	Functional Requirement (Epic)	Sub Requirement (Story/ Sub-Task)
FR-1	User Registration	Registration through Gmail
FR-2	User Confirmation	Confirmation via Email
FR-3	User Communication	Communication can be done through pc or mobile camera.
FR-4	User requirement	Option should be shown for hand signto text and voice conversion and vice versa.
FR-5	Communication requirement	Tutor can be made available to have one to one teaching for user.
FR-6	Regulatory requirements	App shutdown in case of cyber attack
FR-7	Reporting	If any issues found in the application, automatically it will be notified to the developer.
FR-8	Compliance to rules or law	Terms and conditions, private policy, End user subscription agreement.

4.2 Non-Functional Requirements

Following are the non-functional requirements of the proposed solution.

FR.No.	Non-Functional Requirement	Description
NFR-1	**Usability**	The camera captures all expressions including facial expressions and handgestures which can be easily used by all age groups. It can be used by deaf-mute people and their care takers.
NFR-2	**Security**	The system is more secure and information of the customers is also maintained confidentially.
NFR-3	**Reliability**	The system is very liable, it can last for long amounts of time if well maintained.
NFR-4	**Performance**	The performance of the model is efficient. The cost-effective nature of thesystem makes it extremely liable. The latency is very less for the conversion process.
NFR-5	**Availability**	The solution is suitable for different languages and can be used in many countries. It can be trained for all the available sign languages. This model canbe used at any time anywhere.
NFR-6	**Scalability**	The system gives output rapidly. It also predicts quickly when it gets so many inputs at a time. It predicts different typesof sign language at a time. Upto 25000 users can be use this model at a time.

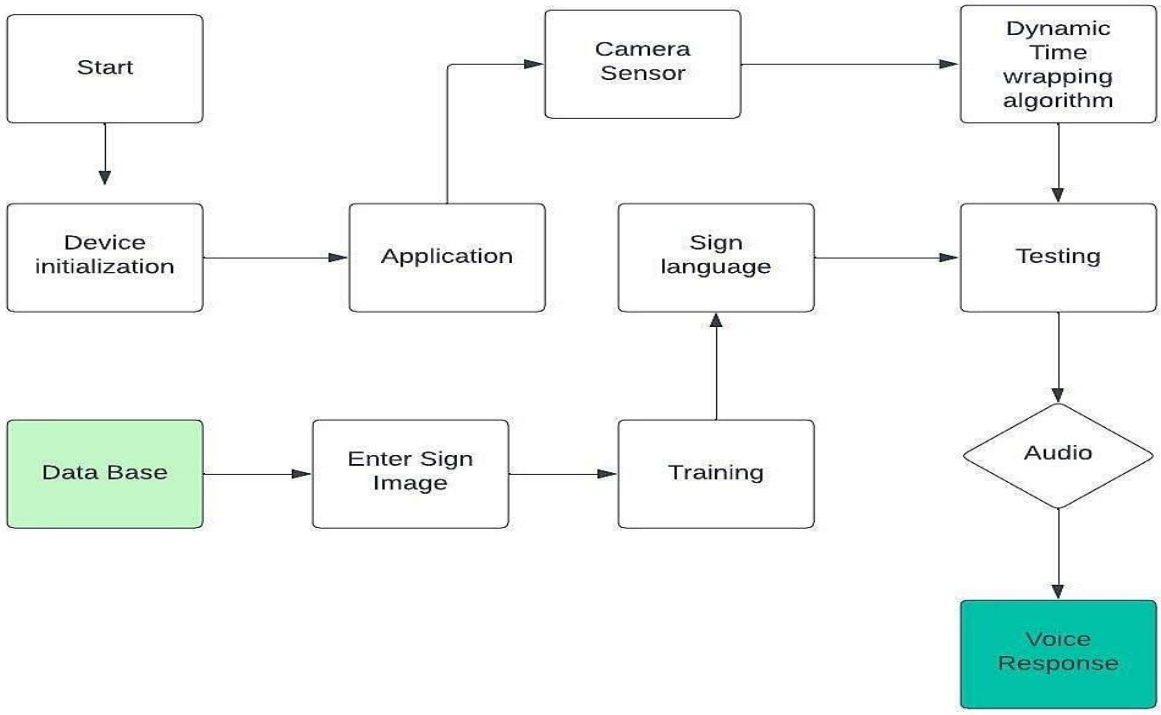
CHAPTER 5

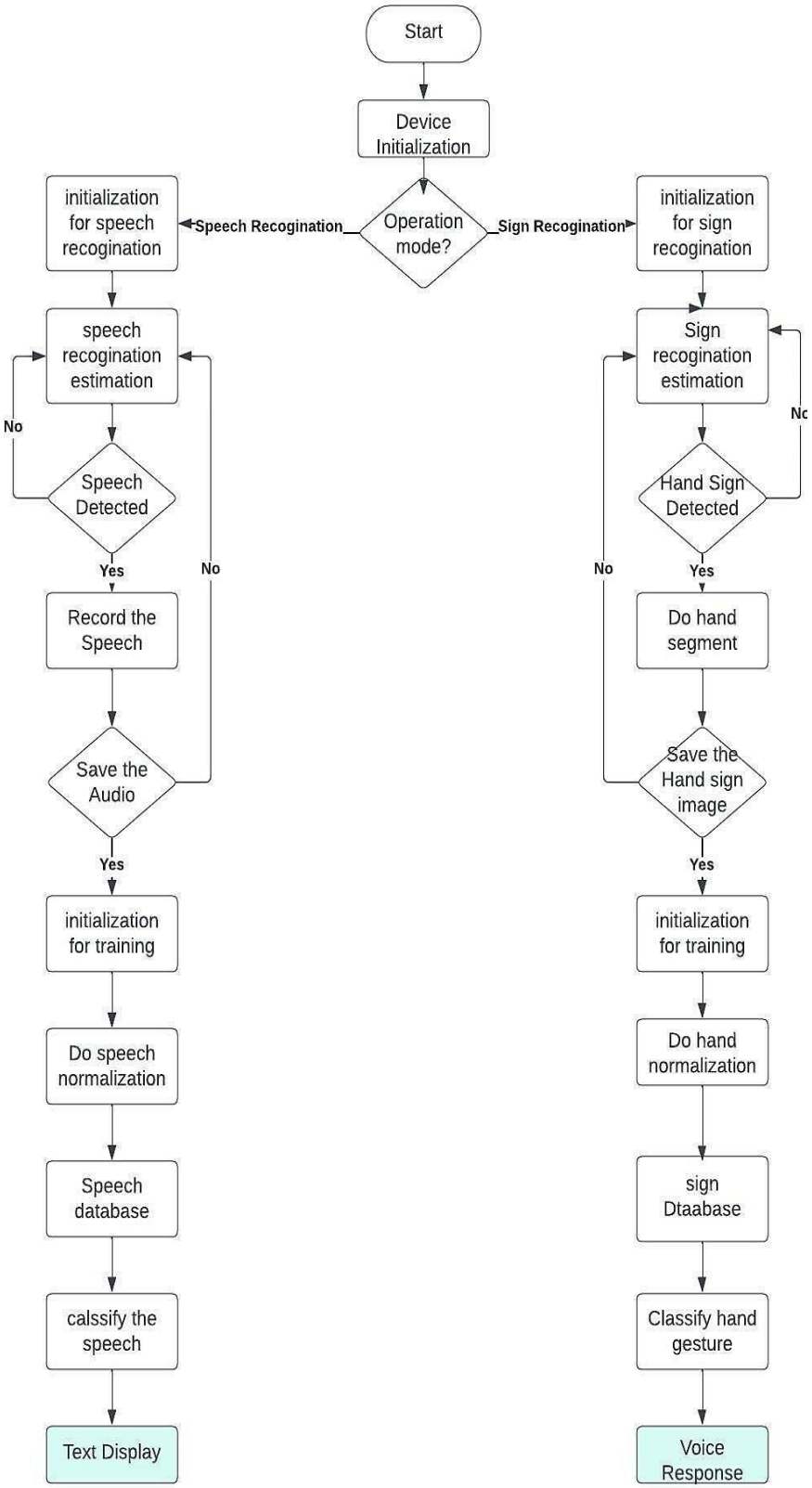
PROJECT DESIGN

5.1 Data Flow Diagrams

A Data Flow Diagram (DFD) is a traditional visual representation of the information flows within a system. A neat and clear DFD can depict the right amount of the system requirement graphically. It shows how data enters andleaves the system, what changes the information, and where data is stored.

Dataflow diagram:




Start
Device Initialization
Operation mode?
Speech Recogination
Sign Recogination
initialization for speech recogination
speech recogination estimation
Speech Detected
No
Yes
Record the Speech
Save the Audio
No
Yes
initialization for training
Do speech normalization
Speech database
calssify the speech
Text Display
initialization for sign recogination
Sign recogination estimation
Hand Sign Detected
No
Yes
Do hand segment
Save the Hand sign image
No
Yes
initialization for training
Do hand normalization
sign Dtaabase
Classify hand gesture
Voice Response

5.2 Solution & Technical Architecture

Solution architecture is a complex process - with many sub-processes - that bridges the gap between business problems and technology solutions. Its goals are to:

- Find the best tech solution to solve existing business problems.
- Describe the structure, characteristics, behavior, and other aspects of the software to project stakeholders.
- Define features, development phases, and solution requirements.
- Provide specifications according to which the solution is defined, managed, and delivered. Solution Architecture Diagram:

Solution Architecture Design:

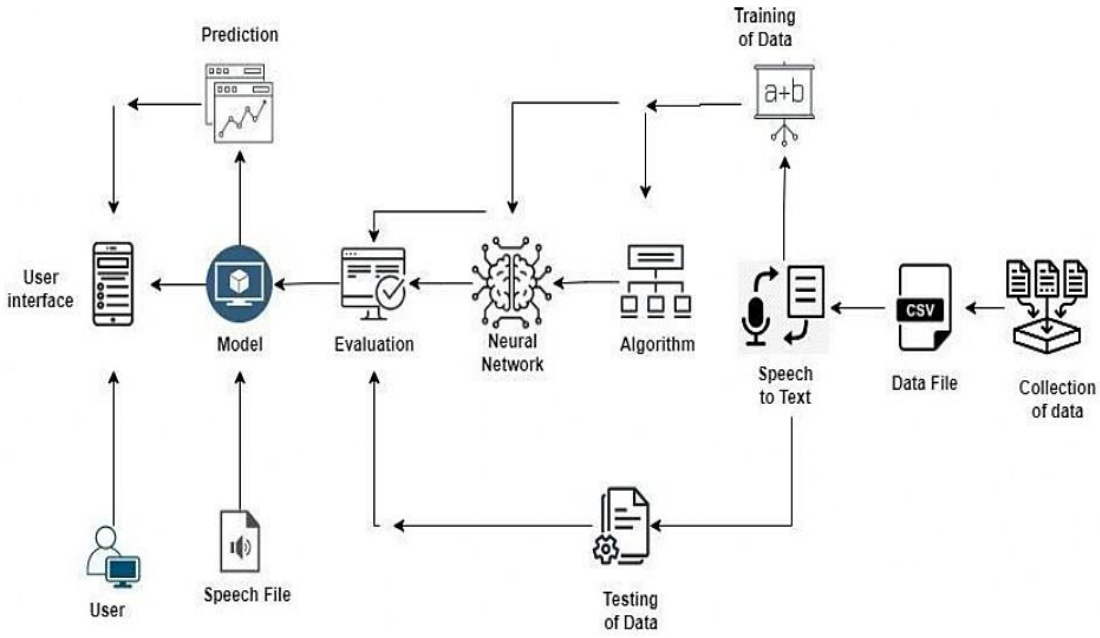


Figure: Architecture design and data flow of the real time communication powered by AI specially abled.

Technical Architecture

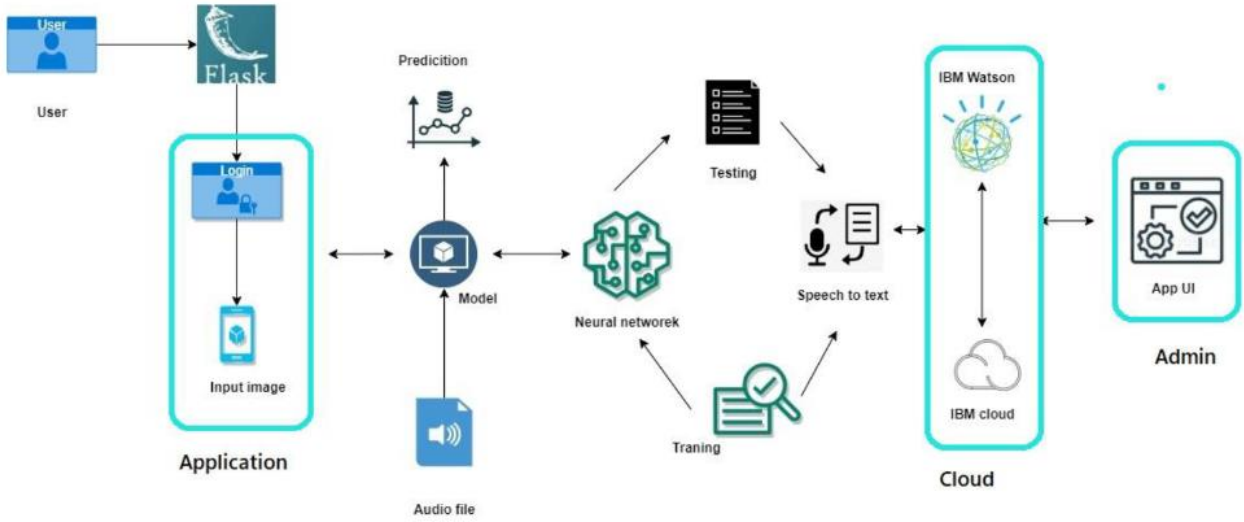


Table-1 : Components & Technologies:

S.No	Component	Description	Technology
1.	User Interface	The way user interacts with application e.g. Web UI,Mobile App, Chatbot etc.	HTML,CSS, JavaScript / Angular JS /React JSetc.
2.	Application Logic-1	It deals with variety of frameworks, libraries and supports required to develop the project	Java / Python
3.	Application Logic-2	Helps in converting hand signs into written Words and it is used to convert text to speech.	IBM WatsonSTT service
4.	Application Logic-3	Provides fast, consistent and accurate answers during the execution phase of the project	IBM Watson Assistant

5.	Database	Images and user inputs are stored and it can be numerical, categorical or time-series data.	MySQL, NoSQL, etc.
6.	Cloud Database	Enables the user to use host database without buying the additional hardware	IBM DB2, IBM Cloud ant etc.
7.	File Storage	File storage should be highly flexible, scalable and effective	IBM Block Storage or Other StorageService or Local Filesystem
8.	External API-1	Used to access the information in the cloud	IBM Weather API, etc.
9.	External API-2	Used to access the information for data driven decision making	Aadhar API,etc.
10.	Machine Learning Model	Machine Learning Model deals with various algorithms that are needed for the implementation	Real time communication using AI for specially abled
11.	Infrastructure (Server/ Cloud)	Application Deployment on Local System / Cloud Local Server Configuration: Install the windows version and execute the installer.	Python flask, Local,Cloud Foundry, Kubernetes, etc.

Table-2: Application Characteristics:

S.No	Characteristics	Description	Technology
1.	Open-Source Frameworks	The frameworks used are	Tensor flow, Theano, RNN, Opensource IBM Watson frameworks.
2.	Security Implementations	The security / access controls implemented, use offirewalls etc.	Certified Watson assistant for encrypted file. Identify, Prevent and respond are followed.
3.	Scalable Architecture	The scalability of architecture (3 – tier, Micro- services)- Data, models, operate at size, speed and complexity	IBM Watson, python, MySQL
4.	Availability	The availability of application - Image and facial recognition, lip reading, text summarization, real time caption available.	IBM Watson
5.	Performance	Design consideration for the performance of Application (number of requests per sec, use of Cache, use of CDN's) etc.	IBM Watson assistant (Full and effective participation, equality of opportunity, accessibility).

5.3 User Stories

Use the below template to list all the user stories for the product.

User Type	Functional Requirement (Epic)	User Story Number	User Story / Task	Acceptance criteria	Priority	Release
Customer *(Desktop user)*	Registration	USN-1	As a user, I can register for the application by entering my email, password, and confirming my password.	I can access my account /dashboard	High	Sprint-1
	Login	USN-2	As a user, I will receive confirmation email once I have registered for the application	I can receive confirmation email & click confirm	High	Sprint-1
	Dashboard	USN-3	As a user, I can select options in dashboard.	I can select options in dashboard.		
Customer *(Desktop user)*	**Main page**	USN-4	As a User, **I can enter the web page once clicked, which provides be the Guidelines** to use the app	I can **enter the web page** once clicked.	Medium	Sprint-1
Customer (Desktop user)	**Guidelines**	USN-5	As a User, **I can give a read through the guidelines to understand the functioning of the app.**	I can **give a read through the guidelines.**	Medium	Sprint-1
Customer *(Desktop user)*	**Convert Sign**	USN-6	As a User, **I can click the button Convert sign**, which directs me towards the Main screen	I can **click the button Convert sign and it direct me to main screen.**	Medium	Sprint-2
Customer *(Desktop user)*	**Camera (Hand movement detection)**	USN-7	As a User, **I can show my hand sign towards the camera which converts the into text manner.**	**I can show my hand sign** towards the camera accurately.	High	Sprint-2
Customer *(Desktop user)*	**Voice mode**	USN-8	Once the text is obtained, as a User **I can click on the voice mode which provides the text in the form of speech.**	**I can click on the voice mode** which provides the text in the form of speech.	High	Sprint-3
Customer Care Executive	**Provide the necessary functionalities required to use the app.**	USN-9	As an Executive, I can **provide the Specifications of Camera required, and other factors** that are required for smooth functioning of the app.	**I can provide the Specifications** of camera required, and other factors	Low	Sprint-1
Customer Care Executive	**Check the performance of the app**	USN-10	As an Executive, I can **check the usage and queries obtained from the end users**.	I can **check the usage and queries** obtained from the end users.	Medium	Sprint-1
Administrator	**Receive queries based on usage**	USN-11	As an Admin, I can **take the queries from the customer care and perform the testing phase again, loading the other signs in the dataset**, in order to make the customers to use the app effectively.	I can **take the queries from the customer care and perform necessary phases again.**	High	Sprint-3

CHAPTER 6

PROJECT PLANNING & SCHEDULING

6.1 Sprint Planning & Estimation

Sprint	Functional Requirement (Epic)	User Story Number	User Story / Task	Story Points	Priority	Team Members
Sprint-1	Registration	USN-1	As a user, I can register for the application by entering my email, password, and confirming my password.	8	High	Rajesh R Rajasekar G
Sprint-2		USN-2	As a user, I will receive confirmation email once I have registered for the application	10	High	Mythrayan S Logavarthan T
Sprint-1	Login	USN-3	As a user, I can log into the application byentering email & password	5	Medium	Rajesh R Mythrayan S
Sprint-2	Dashboard	USN-4	As a user, I can log into my account in a givenDashboard on any time	10	High	Logavarthan T Rajasekar G
Sprint-1	User interface	USN-5	Professional responsible for user requirements &needs	7	High	Rajesh R
Sprint-3	Model initialization	USN-6	Training and testing the model	10	High	Logavarthan T Rajesh R
Sprint-3	Objective	USN-7	Converting hand sign to text and text to speech	10	High	Rajasekar G Mythrayan S
Sprint-4	Privacy	USN-8	The developed application should be secure forthe users	10	High	Rajesh R
Sprint-4	Check the performance of app	USN-9	check the usage andqueries obtained from the end users.	10	Mediun	Mythrayan S

6.2 Sprint Delivery Schedule

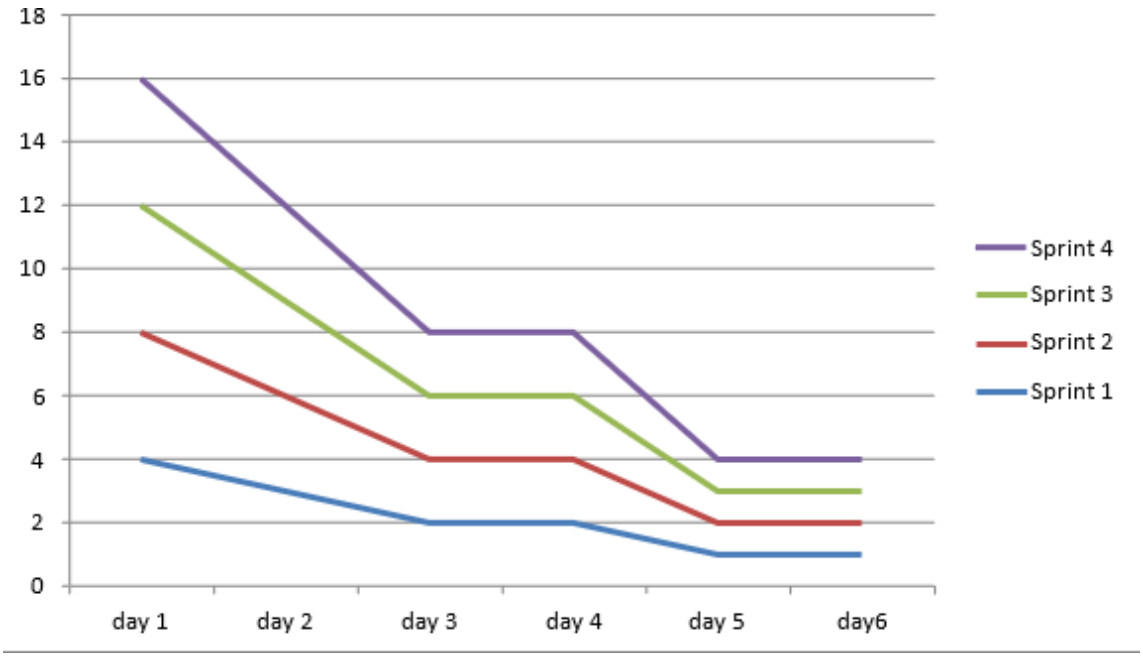
Sprint	Total Story Points	Duration	Sprint Start Date	Sprint End Date (Planned)	Story Points Completed (as on Planned End Date)	Sprint Release Date (Actual)
Sprint-1	20	6 Days	24 Oct 2022	29 Oct 2022	20	29 Oct 2022
Sprint-2	20	6 Days	31 Oct 2022	05 Nov 2022	20	05 Nov 2022
Sprint-3	20	6 Days	07 Nov 2022	12 Nov 2022	20	12 Nov 2022
Sprint-4	20	6 Days	14 Nov 2022	19 Nov 2022	20	19 Nov 2022

Velocity:

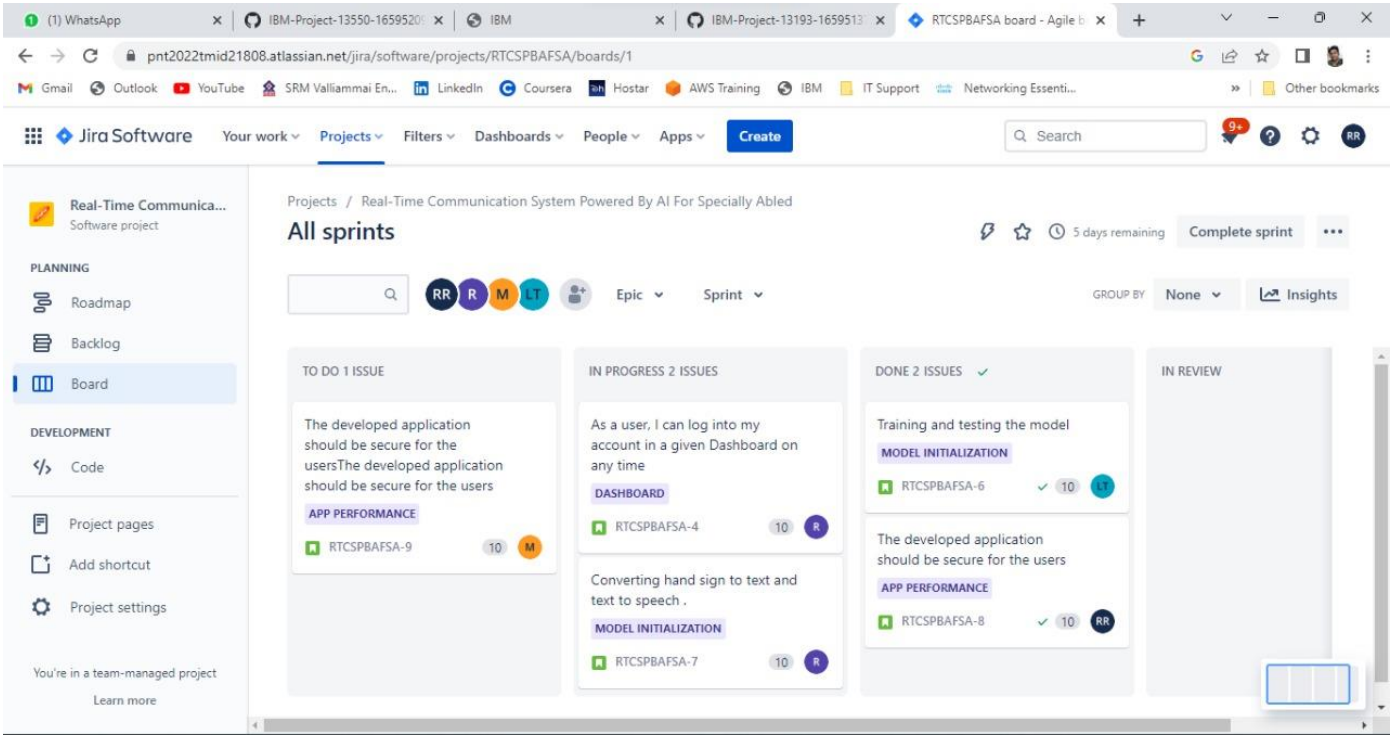
$$AV = \frac{sprint\ duration}{velocity} = \frac{20}{10} = 2$$

Burndown Chart:

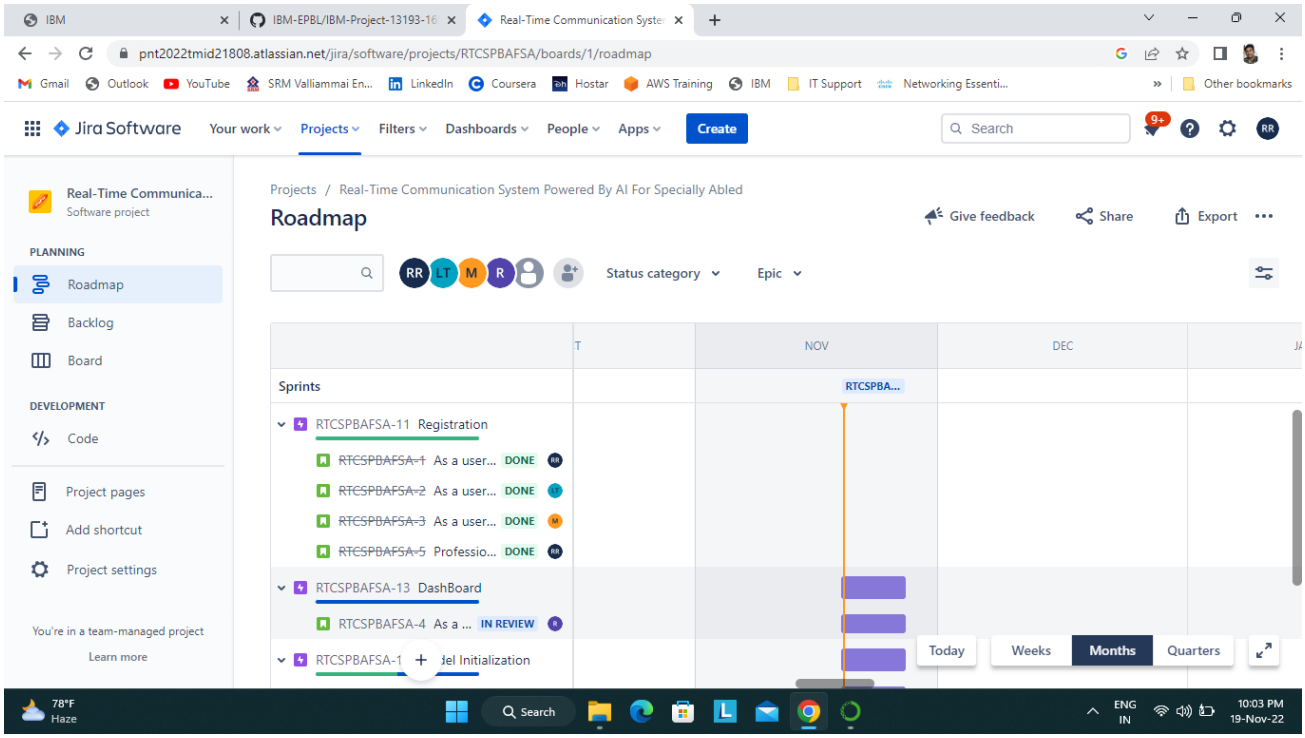
A burn down chart is a graphical representation of work left to do versus time. It is often used in agile software development methodologies such as Scrum. However, burn down charts can be applied to any project containing measurable progress over time.



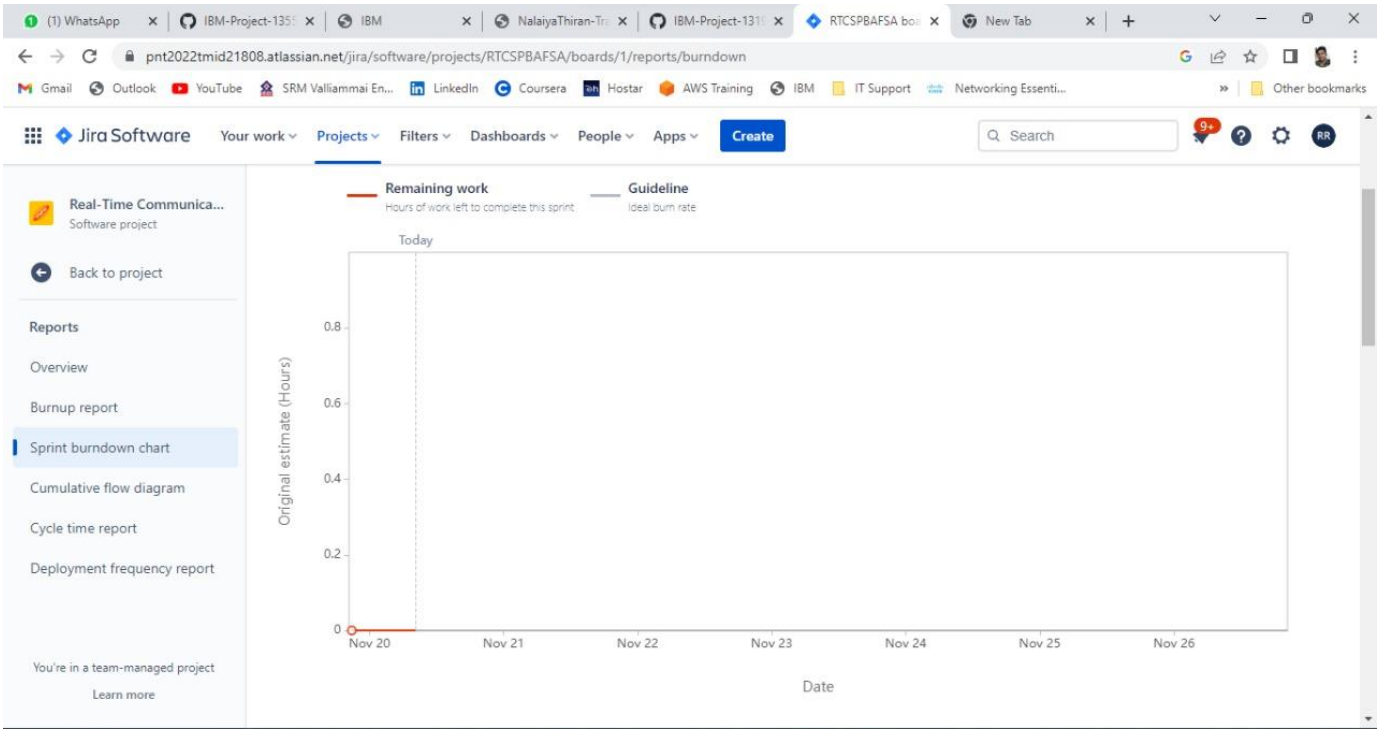
6.3 Reports From Jira



Jira Board



Jira Roadmap – Project Management



Sprint 2 Burndown Chart

CHAPTER 7

CODING AND SOLUTIONING

7.1 Feature 1

The main feature we have included in our model is only the registered user can access the application. In order to increase the security and authorization this feature has been included in the system.

As the first step, the user needs to register in our application. Then the user can able to login for accessing the sign to speech conversion page.

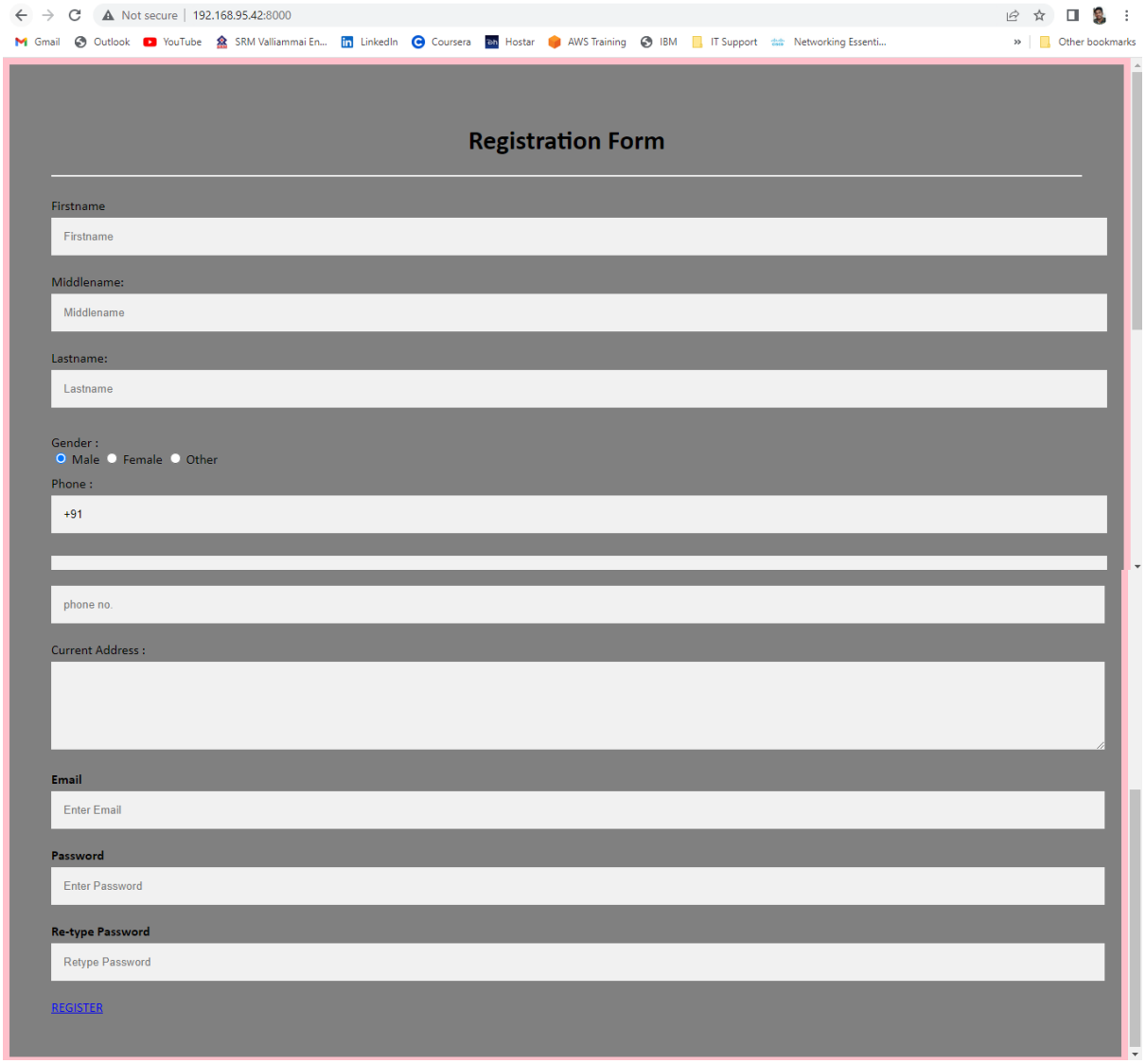
Registration Form (register.html)

```
<!DOCTYPE html>
<html>
<head>
<meta name="viewport" content="width=device-width, initial-scale=1">
<style>
body{
  font-family: Calibri, Helvetica, sans-serif;
  background-color: pink;
}
.container {
    padding: 50px;
  background-color: gray;
}
 input[type=text], input[type=password], textarea {
  width: 100%;
  padding: 15px;
  margin: 5px 0 22px 0;
  display: inline-block;
  border: none;
  background: #f1f1f1;
}
input[type=text]:focus, input[type=password]:focus {
  background-color: orange;
```

```
  outline: none;
}
 div {
          padding: 10px 0;
        }
hr {
  border: 1px solid #f1f1f1;
  margin-bottom: 25px;
}

.registerbtn {
  background-color: #4CAF50;
  color: white;
  padding: 16px 20px;
  margin: 8px 0;
  border: none;
  cursor: pointer;
  width: 100%;
  opacity: 0.9;
}
.registerbtn:hover {
  opacity: 1;
}
</style>
</head>
<body>
<form>
  <div class="container">
  <center>  <h1> Registration Form</h1> </center>
  <hr>

  <label> Firstname </label>
<input type="text" name="firstname" placeholder= "Firstname" size="15" required />
<label> Middlename: </label>
<input type="text" name="middlename" placeholder="Middlename" size="15" required />
<label> Lastname: </label>
<input type="text" name="lastname" placeholder="Lastname" size="15"required />
```

```
<div>
<label>

Gender :
</label><br>
<input type="radio" value="Male" name="gender" checked > Male
<input type="radio" value="Female" name="gender"> Female
<input type="radio" value="Other" name="gender"> Other
 </div>
<label>
Phone :
</label>

<input type="text" name="country code" placeholder="Country Code"  value="+91"
size="2"/>
<input type="text" name="phone" placeholder="phone no." size="10"/ required>
Current Address :
<textarea cols="80" rows="5" placeholder="Current Address" value="address" required>
</textarea>

 <label for="email"><b>Email</b></label>
 <input type="text" placeholder="Enter Email" name="email" required>
     <label for="psw"><b>Password</b></label>

   <input type="password" placeholder="Enter Password" name="psw" required>
     <label for="psw-repeat"><b>Re-type Password</b></label>

   <input type="password" placeholder="Retype Password" name="psw-repeat" required>
   <a href= "/login">REGISTER</a>

</form>
</body>
</html>
```

OUTPUT:

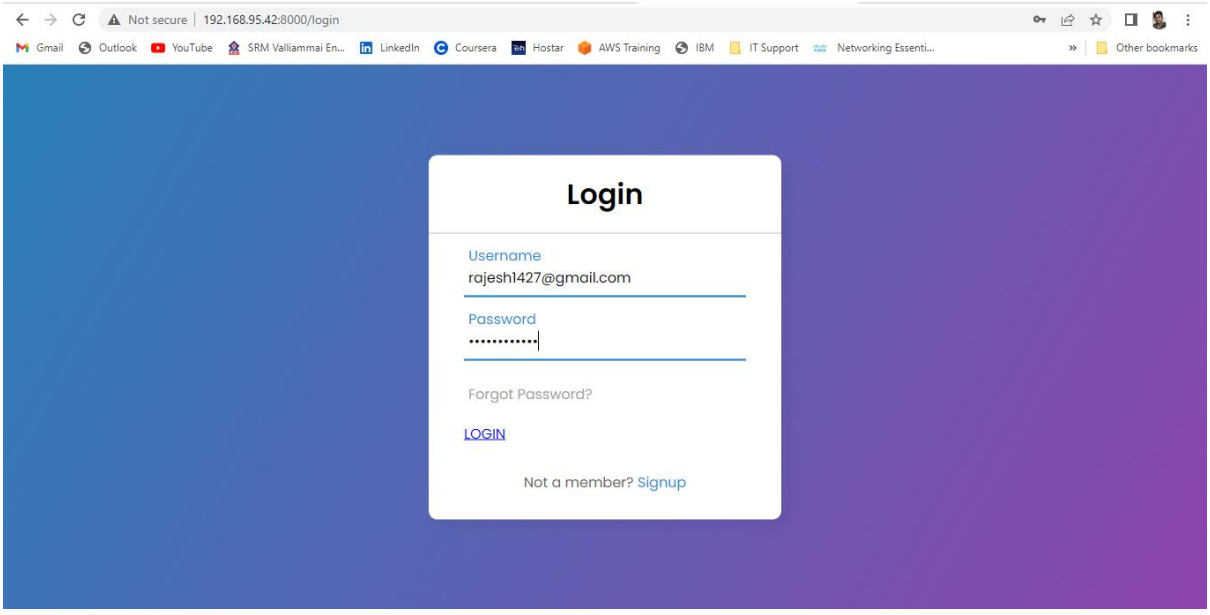


LOGIN PAGE (login.html)

```
<!DOCTYPE html>
<html lang="en" dir="ltr">
  <head>
    <meta charset="utf-8">
    <title>Real-Time Communication System Powered By AI For Specially Abled</title>
    <link rel="stylesheet" type="text/css" href="{{ url_for('static', filename='css/style.css') }}"
  </head>
  <body>
    <div class="center">
```

```
    <h1>Login</h1>
    <form method="post">
      <div class="txt_field">
        <input type="text" required>
        <span></span>
        <label>Username</label>
      </div>
      <div class="txt_field">
        <input type="password" required>
        <span></span>
        <label>Password</label>
      </div>
      <div class="pass">Forgot Password?</div>
      <a href="/index">LOGIN</a>
      <div class="signup_link">
        Not a member? <a href="#">Signup</a>
      </div>
    </form>
  </div>
 </body>
</html>
```

OUTPUT:



7.2 Feature 2

Another feature included in our model is that the sign language prediction on real-time. Through live camera the input is captured, pre-processed and the output is given as speech. Respective alphabet for the hand sign is displayed in the screen and the speech output is also provided for the user.

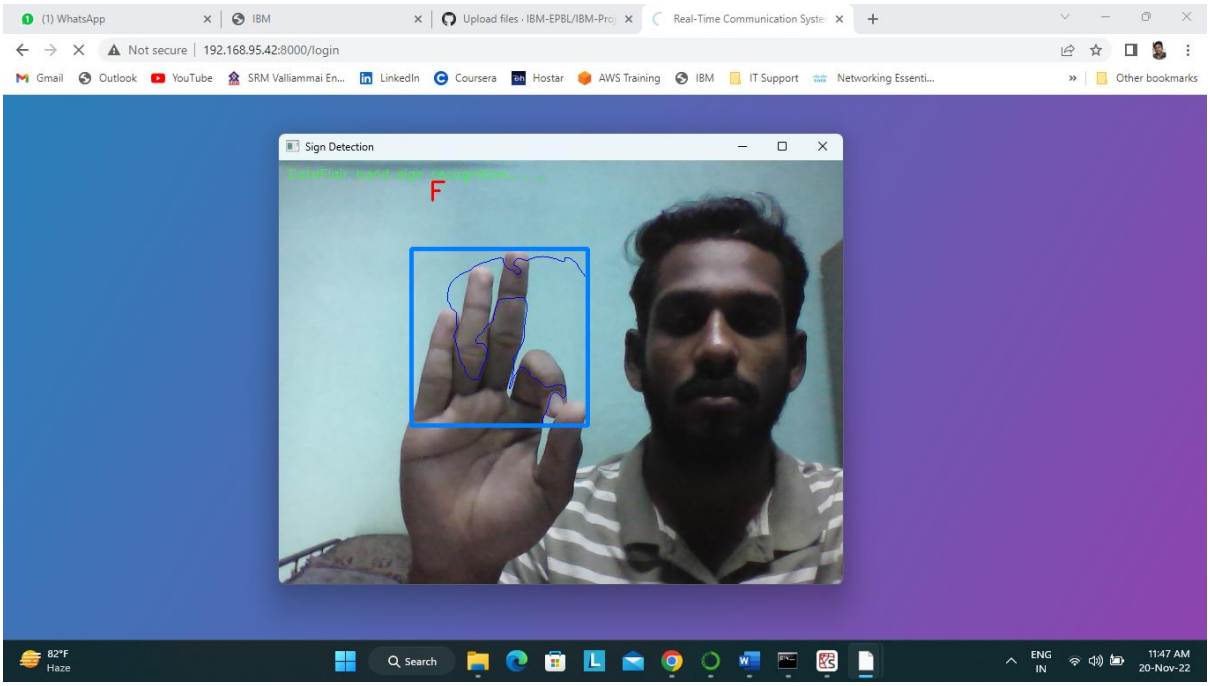
REAL TIME PREDICTION CODE

```
cam = cv2.VideoCapture(0)
num_frames =0
while True:
    ret, frame = cam.read()
    # flipping the frame to prevent inverted image of captured frame...
    frame = cv2.flip(frame, 1)
    frame_copy = frame.copy()
    # ROI from the frame
    roi = frame[ROI_top:ROI_bottom, ROI_right:ROI_left]
    gray_frame = cv2.cvtColor(roi, cv2.COLOR_BGR2GRAY)
    gray_frame = cv2.GaussianBlur(gray_frame, (9, 9), 0)
    if num_frames < 70:
        cal_accum_avg(gray_frame, accumulated_weight)
        cv2.putText(frame_copy, "FETCHING BACKGROUND...PLEASE WAIT",(80, 400),
                    cv2.FONT_HERSHEY_SIMPLEX, 0.9, (0,0,255), 2)
    else:
        # segmenting the hand region
        hand = segment_hand(gray_frame)
        # Checking if we are able to detect the hand...
        if hand is not None:
            thresholded, hand_segment = hand
            # Drawing contours around hand segment
            cv2.drawContours(frame_copy, [hand_segment + (ROI_right,ROI_top)],
                             -1, (255, 0, 0),1)
            cv2.imshow("Thesholded Hand Image", thresholded)
            thresholded = cv2.resize(thresholded, (64, 64))
            thresholded = cv2.cvtColor(thresholded, cv2.COLOR_GRAY2RGB)
            thresholded = np.reshape(thresholded,(1,thresholded.shape[0],thresholded.shape[1],3))
            pred = model.predict(thresholded)
            cv2.putText(frame_copy, word_dict[np.argmax(pred)],(170, 45),
                        cv2.FONT_HERSHEY_SIMPLEX, 1, (0,0,255), 2)

    # Draw ROI on frame_copy
    cv2.rectangle(frame_copy, (ROI_left, ROI_top),
                  (ROI_right,ROI_bottom), (255,128,0), 3)
    # incrementing the number of frames for tracking
    num_frames += 1
```

```
    # Display the frame with segmented hand
    cv2.putText(frame_copy, "Hand sign recognition",
    (10, 20), cv2.FONT_ITALIC, 0.5, (51,255,51), 1)
    cv2.imshow("Sign Detection", frame_copy)
    # Close windows with Esc
    k = cv2.waitKey(1) & 0xFF
    if k == 'q':
        break

# Release the camera and destroy all the windows
cam.release()
cv2.destroyAllWindows()
```

OUTPUT:

The Predicted alphabet is “F”. Accuracy of the model is 98%.



CHAPTER 8

TESTING

8.1 Test Cases

This report shows the number of test cases that have passed, failed, and untested by our model.

Test case ID	Feature Type	Component	Test Scenario	Pre-Requisite	Steps To Execute	Test Data	Expected Result	Actual Result	Status	Commnets	TC for Automation(Y/N	BUG ID	Executed By
Detection_01	User Interface Webpage	Registration Page	Verify whether the user is able to view the Registration Page	Active server with Internet Connectivity and Frontend Code	1. Enter Website URL and Search the URL 2. Display the Register Page to the user	http://192.168.95.42:8000/	Register Page will be display withthe Process of Front end	Working as expected	Pass	User can view the registration page	YES	-	PNT2022TNID21808
Detection_02	User Interface Webpage	Registration Page	Verify whether the user is able to register with credentials	Active server with Internet Connectivity and Frontend Code	1. Enter Website URL and click go 2. Display the Register Page to the User 3. Able to register in registration page	http://192.168.95.42:8000/	Successful Registration	Working as expected	Pass	User can register	YES	-	PNT2022TNID21808
Detection_03	User Interface Webpage	Login page	Verify whether the user to view Login Page	Active server with Internet Connectivity and Frontend Code, HTML Search tag	1.Enter URL and click go 2. Display the Login Page to the User	http://192.168.95.42:8000/	Login Page will be display with the Process of Front end	Working as expected	Pass	User can view the login page	YES	-	PNT2022TNID21808
Detection_04	User Interface Webpage	Login page	Verify user is able to log into application with Valid credentials	Active server with Internet Connectivity and Frontend Code, HTML Search tag with valid URL	1.Enter URL and click go 2. Display the Login Page to the User 3. Able to login in Login Page	http://192.168.95.42:8000/	Successful Login.	Working as expected	Pass	User can login into web app	YES	-	PNT2022TNID21808
Detection_05	Conversion	Detection Page	Verify user is able to see the Conversion Page	Active server with Internet Connectivity	1.Enter URL and click go 2. Redirect to Conversion Page	http://192.168.95.42:8000/	Able to see Conversion page	Working as expected	Pass	User can see the conversion page	YES	-	PNT2022TNID21808
Detection_06	Prediction	Detection Page	Verify user can able to predict hand sign through live video Feed	Active server with Internet Connectivity	1.Enter URL and click go 2.Redirect to Conversion Page 3. View Hand Sign in Live Camera	http://192.168.95.42:8000/	Able to show hand sign in camera	Working as expected	Pass	User can show hands in front of camera	YES	-	PNT2022TNID21808
Detection_07	Prediction	Detection Page	Check the user is abble to get the corressponding Alphabet	Active server with Internet Connectivity	1.Enter URL and click go 2.Redirect to Conversion Page 3. View Hand Sign in Live Camera 4. Get Coressponding Alphabet for Hand Sign given	http://192.168.95.42:8000/	Able to get the predicted alphabet for the hand sign	Working as expected	Pass	User can get the conversion results.	YES	-	PNT2022TNID21808

8.2 User Acceptance Testing

Purpose of Document

The purpose of this document is to briefly explain the test coverage and open issues of the Real time communication powered by AI for specially abled project at the time of the release to User Acceptance Testing (UAT). 2

.

Defect Analysis

This report shows the number of resolved or closed bugs at each severity level, and how they were resolved by our Model Sign language site).

Resolution	Severity 1	Severity 2	Severity 3	Severity 4	Subtotal
By Design	6	4	2	2	14
Duplicate	0	1	1	1	3
External	4	1	0	5	10
Fixed	9	6	5	14	34
Not Reproduced	1	1	0	1	3
Skipped	1	0	1	1	3
Won't Fix	0	0	2	2	4
Totals	21	13	11	26	71

Test Case Analysis

This report shows the number of test cases that have passed, failed, and untested by our model.

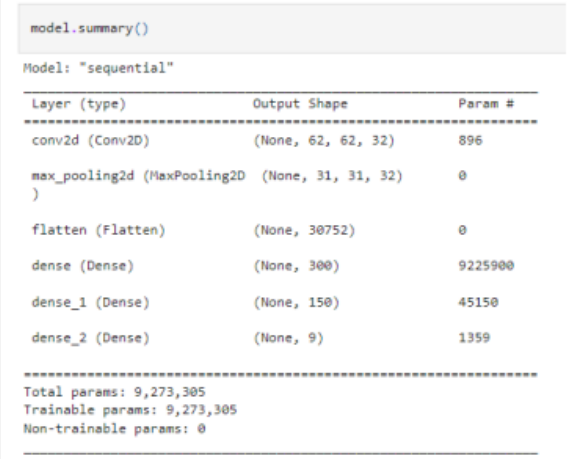
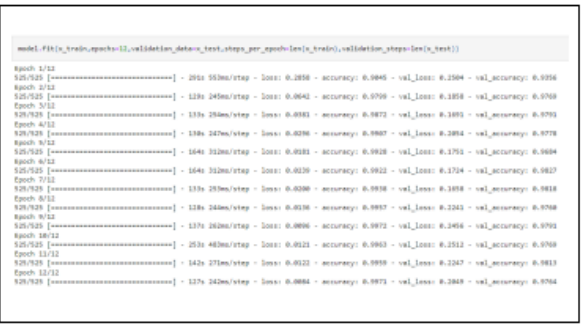
Section	Total Cases	Not Tested	Fail	Pass
Register	20	0	4	16
Login	25	0	6	19
Redirect to conversion page	40	0	8	32
Real Time Video Analysis	60	0	10	50
Hand sign prediction	70	0	23	47
Final Model Output	70	0	54	16
Security	25	0	4	21

CHAPTER 9

RESULTS

9.1 Performance Metrics

Project team shall fill the following information in model performance testing template.

S.No.	Parameter	Values	Screenshot
1.	Model Summary	Model: "sequential" _____ __ Layer (type) Output Shape Param # ========== ==conv2d (Conv2D) (None, 62, 62, 32) 896 max_pooling2d (None, 31, 31, 32) 0 flatten (Flatten) (None, 30752) 0 dense (Dense) (None, 300) 9225900 dense_1 (Dense) (None, 150) 45150 dense 2 (Dense) (None, 9) 1359 ========== == Total params: 9,273,305 Trainable params: 9,273,305 Non-trainable params: 0	
2.	Accuracy	Training Accuracy - 0.9971 Validation Accuracy - 0.9764	

MODEL SUMMARY:

- Each layer has an output and its shape is shown in the "Output Shape" column. Each layer's output becomes the input for the subsequent layer.
- The "Param #" column shows you the number of parameters that are trained for each layer.

- The total number of parameters is shown at the end, which is equal to the number of trainable and non-trainable parameters. In this model, all the layers are trainable.

Accuracy:

- "loss" refers to the loss value over the training data after each epoch. This is what the optimization process is trying to minimize with the training so, the lower, the better.
- "accuracy" refers to the ratio between correct predictions and the total number of predictions in the training data. The higher, the better. This is normally inversely correlated with the loss.

Model summary image:

```
In [13]: model.summary()
```

```
Model: "sequential"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 conv2d (Conv2D)             (None, 62, 62, 32)        896

 max_pooling2d (MaxPooling2D  (None, 31, 31, 32)       0
 )

 flatten (Flatten)           (None, 30752)             0

 dense (Dense)               (None, 300)               9225900

 dense_1 (Dense)             (None, 150)               45150

 dense_2 (Dense)             (None, 9)                 1359

=================================================================
Total params: 9,273,305
Trainable params: 9,273,305
Non-trainable params: 0
_________________________________________________________________
```

Accuracy image:

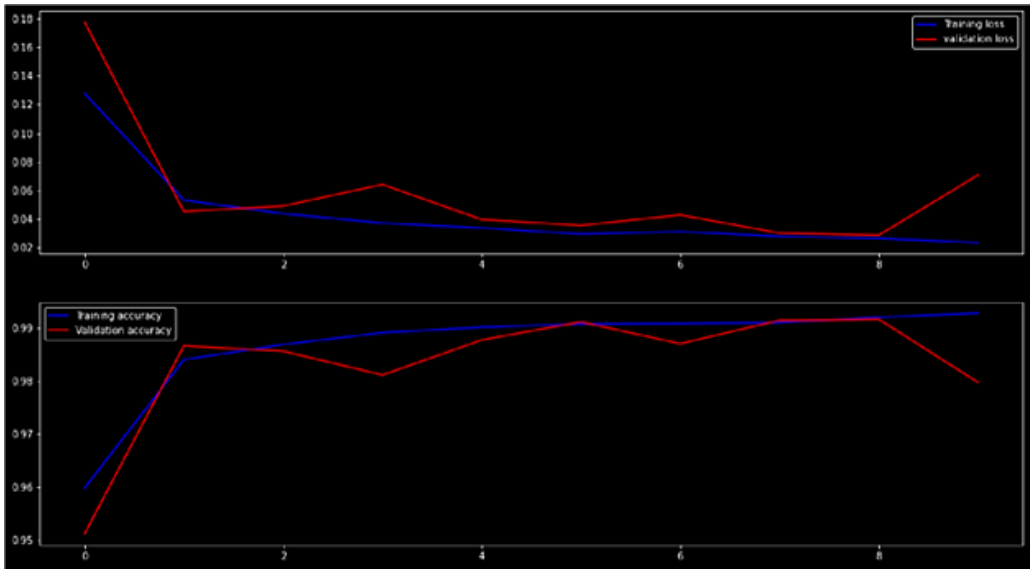
```
In [18]: model.fit(x_train,epochs=12,validation_data=x_test,steps_per_epoch=len(x_train),validation_steps=len(x_test
```

```
Epoch 1/12
525/525 [==============================] - 291s 553ms/step - loss: 0.2858 - accuracy: 0.9045 - val_los
s: 0.2504 - val_accuracy: 0.9356
Epoch 2/12
525/525 [==============================] - 129s 245ms/step - loss: 0.0642 - accuracy: 0.9799 - val_los
s: 0.1858 - val_accuracy: 0.9769
Epoch 3/12
525/525 [==============================] - 133s 254ms/step - loss: 0.0381 - accuracy: 0.9872 - val_los
s: 0.1691 - val_accuracy: 0.9791
Epoch 4/12
525/525 [==============================] - 130s 247ms/step - loss: 0.0296 - accuracy: 0.9907 - val_los
s: 0.2054 - val_accuracy: 0.9778
Epoch 5/12
525/525 [==============================] - 164s 312ms/step - loss: 0.0181 - accuracy: 0.9928 - val_los
s: 0.1751 - val_accuracy: 0.9684
Epoch 6/12
525/525 [==============================] - 164s 312ms/step - loss: 0.0239 - accuracy: 0.9922 - val_los
s: 0.1724 - val_accuracy: 0.9827
Epoch 7/12
525/525 [==============================] - 133s 253ms/step - loss: 0.0200 - accuracy: 0.9938 - val_los
s: 0.1658 - val_accuracy: 0.9818
Epoch 8/12
525/525 [==============================] - 128s 244ms/step - loss: 0.0136 - accuracy: 0.9957 - val_los
s: 0.2241 - val_accuracy: 0.9760
Epoch 9/12
525/525 [==============================] - 137s 262ms/step - loss: 0.0096 - accuracy: 0.9972 - val_los
s: 0.2456 - val_accuracy: 0.9791
Epoch 10/12
525/525 [==============================] - 253s 483ms/step - loss: 0.0121 - accuracy: 0.9963 - val_los
s: 0.2512 - val_accuracy: 0.9769
Epoch 11/12
525/525 [==============================] - 142s 271ms/step - loss: 0.0122 - accuracy: 0.9959 - val_los
s: 0.2247 - val_accuracy: 0.9813
Epoch 12/12
```

```
525/525 [==============================] - 127s 242ms/step - loss: 0.0084 - accuracy: 0.9971 - val_los
s: 0.2049 - val_accuracy: 0.9764
```

```
Out[18]: <keras.callbacks.History at 0x1c7e7ed93a0>
```

ACCURACY TESTING:



CHAPTER 10

ADVANTAGES & DISADVANTAGES

Advantages

The major merits of our real-time communication system are as follows,

- Our system brings many benefits to all people regardless of whether they are deaf or struggling with their hearing. As well as **helping people to communicate and fully express themselves**.
- It also improves social skills by increasing their confidence and self-esteem of person with disabilities in communication.
- It is a User-friendly application and can be easily access by all the age group people.
- Uneducated people also can be used it easily. No need of any prior knowledge before using this application.
- Bridges the gap between Deaf/Dump people with normal people.
- Implementation is easier and color parameters can also include in the system.
- Model Accuracy is 98%.

Disadvantages

The major demerits of our real-time communication system are as follows,

- Network issues may delay to access the feature.
- Dynamic signs are difficult to include.
- Model Training requires huge amount of data.
- Pre-processing of Images need to be done. Hence require high storage capacity.
- Improper sign language will lead the people to misunderstood.
- Need good internet connectivity.

11. CONCLUSION

Nowadays, applications need several kinds of images as sources of information for elucidation and analysis. Several features are to be extracted so as to perform various applications. When an image is transformed from one form to another such as digitizing, scanning, and communicating, storing, etc. degradation occurs. Therefore, the output image has to undertake a process called image enhancement, which contains of a group of methods that seek to develop the visual presence of an image. The Proposed system introduces an efficient and fast algorithm for identification of the number of fingers opened in a gesture representing an alphabet of the Sign Language. Machine learning and artificial intelligence concepts to take visual inputs of sign language's hand gestures and generate easily recognizable form of outputs. Thus it used to develop an intelligent system which can act as a translator between the sign language and the alphabet conversion.

12. FUTURE SCOPE

The system can be useful for static ISL numeral signs only. The ISL recognizer system cannot be considered as a complete system, as for complete recognition of sign language, we have to include ISL alphabets, words and sentences. These signs can be included in future. Also, other feature extraction algorithms like Wavelet transform, Invariant moments, Shape lets descriptors and other existing methods can be included in conducting experiments for improvement in the results. Other classifiers like multi class Support Vector Machine (SVM), Principal Component Analysis (PCA) and Linear Discriminant Analysis (LDA) or a combination of these classifiers can be included in conducting experiments to improve the recognition rate.

13.APPENDIX

13.1 Source Code:

app.py

```
from flask import Flask,render_template
import cv2
from keras.models import load_model
import numpy as np
import pyttsx3 #to convert text to speech
from skimage.transform import resize
app = Flask(__name__,template_folder="templates")
# Loading the model
model=load_model("Model/sgnlan.h5")
print("Loaded model from disk")

background = None
accumulated_weight = 0.5
ROI_top = 100
ROI_bottom = 300
ROI_right = 150
ROI_left = 350

word_dict = { 0:'A', 1:'B',  2:'C', 3: 'D', 4:'E', 5:'F', 6:'G',7:'H', 8:'I' }
vals = ['A', 'B','C','D','E','F','G','H','I']

@app.route('/', methods=['GET'])
def index():
    return render_template('register.html')

@app.route('/login', methods=['GET'])
def home():
    return render_template('login.html')


@app.route('/upload', methods=['GET', 'POST'])
def predict():
    def cal_accum_avg(frame, accumulated_weight):

        global background

        if background is None:
            background = frame.copy().astype("float")
            return None

        cv2.accumulateWeighted(frame, background, accumulated_weight)

    def segment_hand(frame, threshold=25):
        global background
        diff = cv2.absdiff(background.astype("uint8"), frame)
        _ , thresholded = cv2.threshold(diff, threshold,255,cv2.THRESH_BINARY)
        #Fetching contours in the frame (These contours can be of hand or any other object in foreground) …
        contours, hierarchy =cv2.findContours(thresholded.copy(), cv2.RETR_EXTERNAL,cv2.CHAIN_APPROX_SIMPLE)
        # If length of contours list = 0, means we didn't get any contours...
        if len(contours) == 0:
            return None
```

```
        else:
            # The largest external contour should be the hand
            hand_segment_max_cont = max(contours, key=cv2.contourArea)

            # Returning the hand segment(max contour) and the thresholded image of hand...
            return (thresholded, hand_segment_max_cont)

    cam = cv2.VideoCapture(0)
    num_frames =0
    while True:
        ret, frame = cam.read()
        # flipping the frame to prevent inverted image of captured frame...
        frame = cv2.flip(frame, 1)
        frame_copy = frame.copy()
        # ROI from the frame
        roi = frame[ROI_top:ROI_bottom, ROI_right:ROI_left]
        gray_frame = cv2.cvtColor(roi, cv2.COLOR_BGR2GRAY)
        gray_frame = cv2.GaussianBlur(gray_frame, (9, 9), 0)
        if num_frames < 70:
            cal_accum_avg(gray_frame, accumulated_weight)
            cv2.putText(frame_copy, "FETCHING BACKGROUND...PLEASE WAIT",(80, 400),
                        cv2.FONT_HERSHEY_SIMPLEX, 0.9, (0,0,255), 2)
        else:
            # segmenting the hand region
            hand = segment_hand(gray_frame)
            # Checking if we are able to detect the hand...
            if hand is not None:
                thresholded, hand_segment = hand
                # Drawing contours around hand segment
                cv2.drawContours(frame_copy, [hand_segment + (ROI_right,ROI_top)],
                                 -1, (255, 0, 0),1)
                cv2.imshow("Thesholded Hand Image", thresholded)
                thresholded = cv2.resize(thresholded, (64, 64))
                thresholded = cv2.cvtColor(thresholded, cv2.COLOR_GRAY2RGB)
                thresholded = np.reshape(thresholded,(1,thresholded.shape[0],thresholded.shape[1],3))
                pred = model.predict(thresholded)
                result =word_dict[np.argmax(pred)]
                cv2.putText(frame_copy, result,(170, 45),
                            cv2.FONT_HERSHEY_SIMPLEX, 1, (0,0,255), 2)
                text_speech =  pyttsx3.init()
                rate=100
                text_speech.setProperty("rate", rate)
                text_speech.say (result)
                text_speech.runAndWait()

            # Draw ROI on frame_copy
            cv2.rectangle(frame_copy, (ROI_left, ROI_top),
                          (ROI_right,ROI_bottom), (255,128,0), 3)
            # incrementing the number of frames for tracking
            num_frames += 1
            # Display the frame with segmented hand
            cv2.putText(frame_copy, "DataFlair hand sign recognition_ _ _",
            (10, 20), cv2.FONT_ITALIC, 0.5, (51,255,51), 1)
            cv2.imshow("Sign Detection", frame_copy)
            # Close windows by pressing the key 'q'
            if cv2.waitKey(0) & 0xFF == ord('q'):
                break
```

```
108
109         # Release the camera and destroy all the windows
110         cam.release()
111         cv2.destroyAllWindows()
112
113
114         return render_template("index.html")
115
116
117     if __name__ == '__main__':
118           app.run(host='0.0.0.0', port=8000, debug=False)
```

13.2 GitHub & Project Demo Link:

GitHub Link:

- **https://github.com/IBM-EPBL/IBM-Project-13193-1659513703**

Project Demo Link:

- **https://drive.google.com/drive/folders/15NC5kQDCivPWqU491nFCnU1H0JZ3usde?usp=sharing**