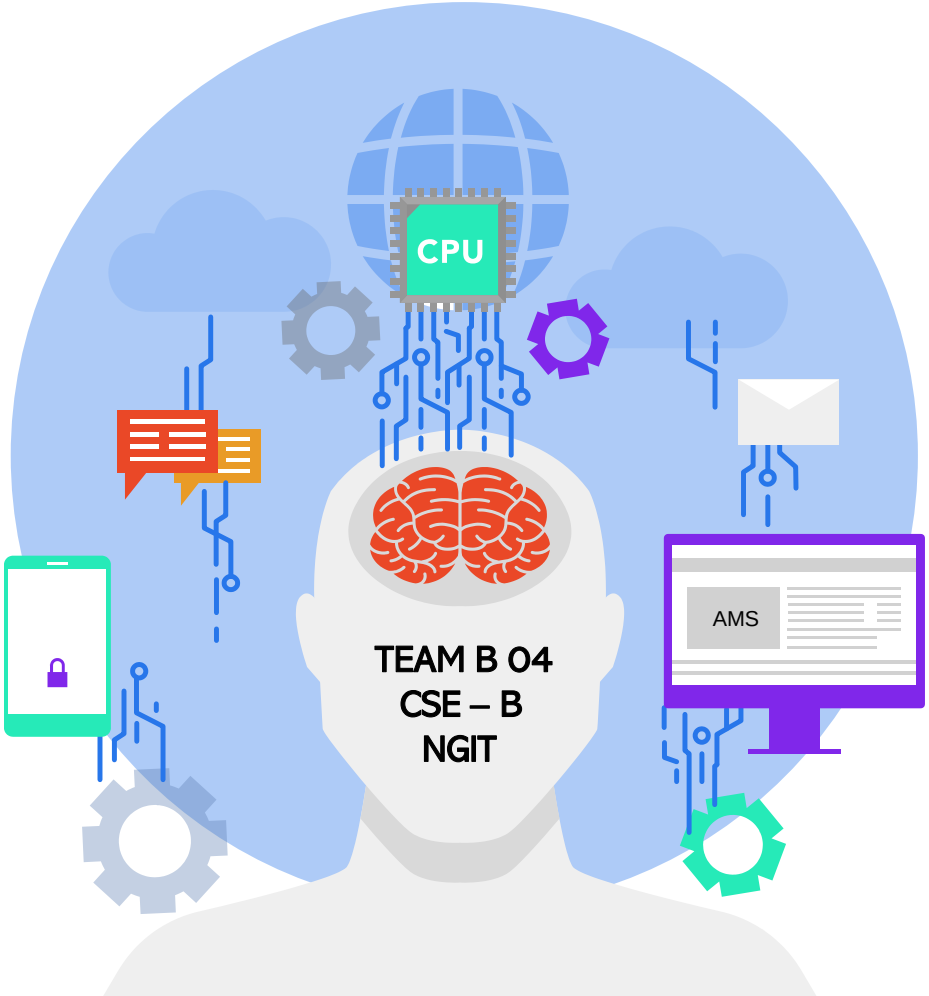


GESTURE RECOGNITION BASED ON COMPUTER-VISION AND MEDIAPIPE

NEIL GOGTE INSTITUTE OF TECHNOLOGY

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

GESTURE RECOGNITION BASED ON COMPUTER-VISION AND MEDIAPIPE

INTERNAL GUIDE :

Dr. B. Krishna, Professor, CSE

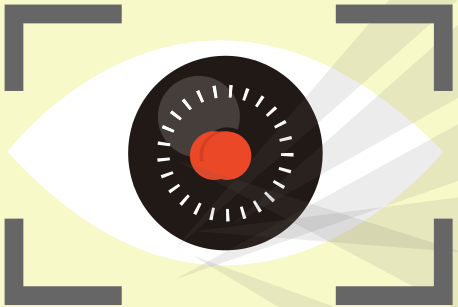
PRESENTED BY

Abhishek Choudhary 2453-18-733-121

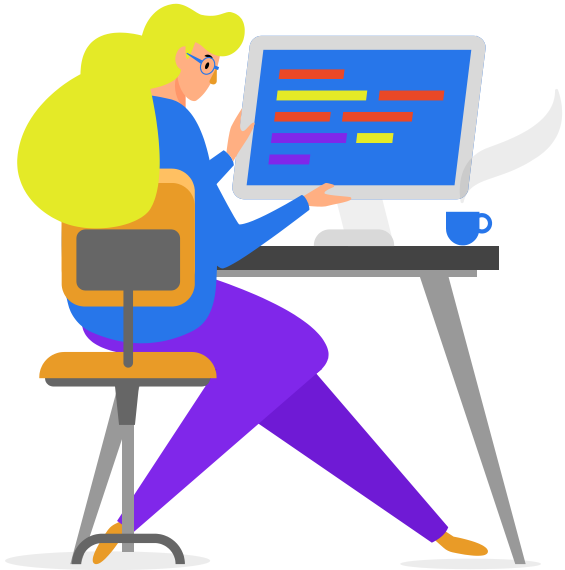
Gangishetty Mounik Teja 2453-18-733-080

Siddartha Pullakhandam 2453-18-733-107

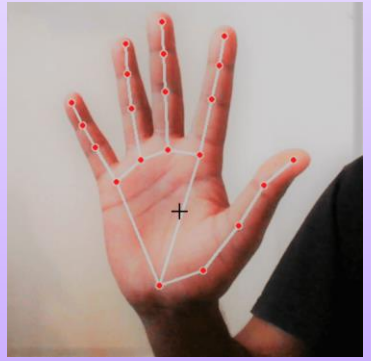
CONTENTS



INTRODUCTION



- In this project we try to recognize different hand gestures that are used.

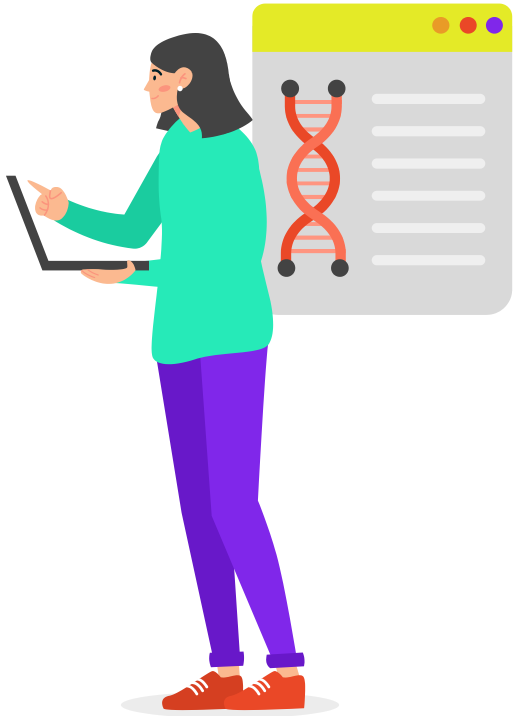


- By using a web-camera in real time, we recognize the hand-gestures by using computer vision, trained machine learning models and different libraries and frameworks.
- We picked the **k-nearest neighbors algorithm (k-NN)** as the classifier. It's simple and easy to start with.
- The algorithm determines the object's class based on the closest samples in the training set.

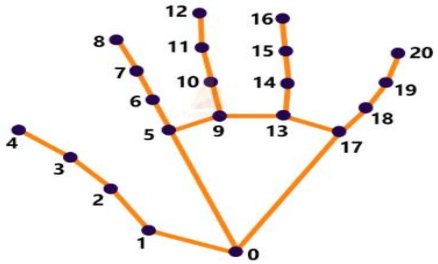
EXISTING SYSTEM

- There have been a few developments on gesture language recognition with the usage of computer vision and image processing techniques. The overall purpose of these developments is to provide an alternative method to aid the disabled people in terms of communicating with each other to deliver a message or to depict an emotion without the usage of human voice or emotions.

- A model of Hand Gesture recognition made use of an Artificial Neural Network in order to train and recognize a static hand gesture and Jiahui Wu had equipped an acceleration-based gesture recognition approach, called FDSVM (Frame- based Descriptor and multi-class SVM), which needs only a wearable 3-dimensional accelerometer from which the gesture is collected and represented in the form of a frame-based descriptor."



PROPOSED SYSTEM



POINTS	NAME
0	WRIST
1	THUMB CMC
2	THUMB MCP
3	THUMB IP
4	THUMB TIP
5	INDEX FINGER MCP
6	INDEX FINGER PIP
7	INDEX FINGER DIP
8	INDEX FINGER TIP
9	MIDDLE_FINGER_MCP
10	MIDDLE FINGER PIP
11	MIDDLE FINGER DIP
12	MIDDLE FINGER TIP
13	RING FINGER MCP
14	RING FINGER PIP
15	RING FINGER DIP
16	RING FINGER TIP
17	PINKY MCP

- We are able to detect and train the hand gestures using the mediapipe API. Media pipe API is an ML pipeline that provides services of palm detection and drawing landmarks on the hand.
- The API first recognizes the palm in the frame and then draws landmarks in the hand. Following palm detection over the entire image, the hand landmark model in the media pipe API uses regression to accomplish exact key point localization of 21 3D hand-knuckle coordinates within the detected hand regions, i.e. direct coordinate prediction.
- Even with partially visible hands and self-occlusions, the model develops a consistent internal hand posture representation.
- These landmarks are used for the prediction of the name of the gestures. The trained model has an accuracy of 95.7% in comparison to the existing models.
- This is a real-time prediction, as we change our hand gestures the name appearing in the display also changes. The model continuously predicts the name of each gesture that is found in each of the frames.

OPERATING ENVIRONMENT

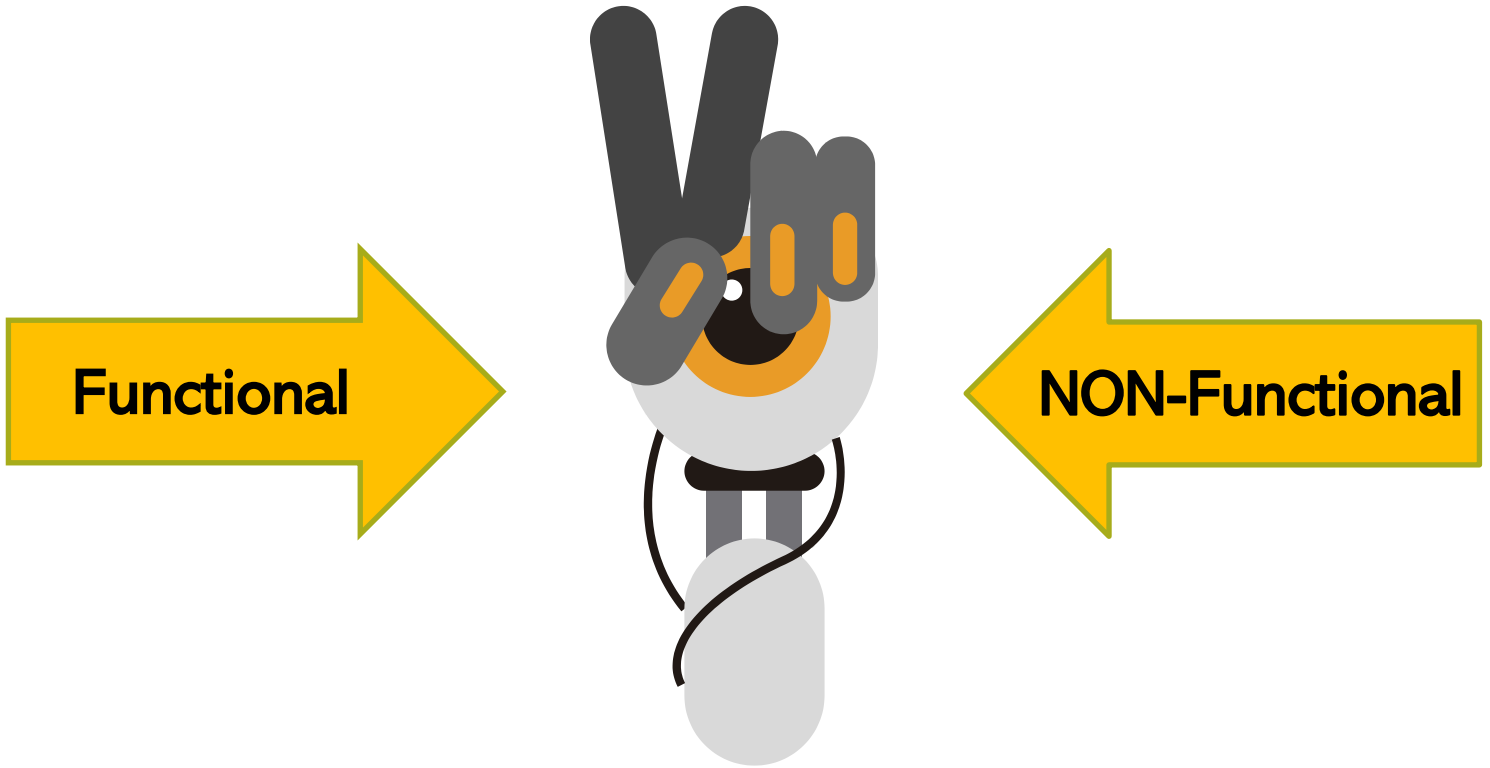
Software requirements:

- Operating System : Any updated Operating System
- Language used : Python
- Libraries : OpenCV, Tensorflow, NumPy, MediaPipe
- Development Kit : Any python supported IDE

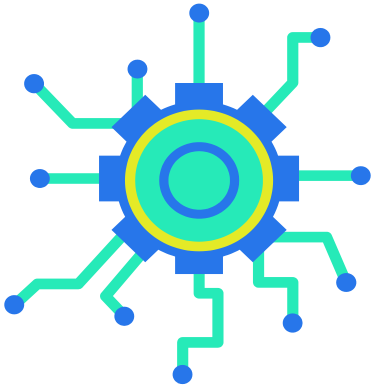
Hardware requirements:

- Processor : Intel Pentium® Dual Core Processor (Min)
- Speed : 2.9 GHz (Min)
- RAM : 4 GB (Min)
- Hard Disk : 2 GB (Min)

REQUIREMENTS



FUNCTIONAL REQUIREMENTS



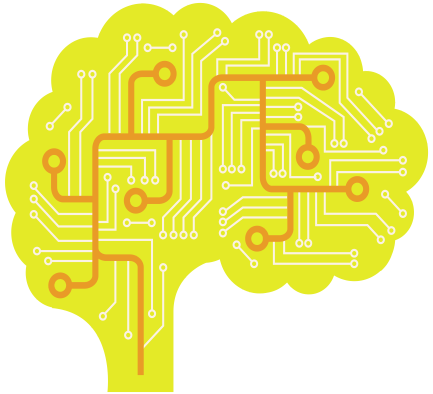
- In this model hands are the main input sources that are used to give test cases to the trained Machine Learning model and this model uses features like landmarks to complete the project and give the output.

- Landmark is the identification portion in the model, landmarks are given at different points on palm of a hand using media pipe library, we have a list of points that identify different parts of the palm, and then gestures are made.

- In this project, we try to use computer vision, machine learning and try to recognize hand gestures. The process starts with a webcam feed which captures the input of hand movement frame by frame and each frame is used in recognition of gestures in real-time this is done when the palm is assigned landmarks which are then given to the trained Machine Learning model which tests the gesture given and then recognizes the given gesture using the landmarks, position of landmarks in the plane.

- Gestures are made using hands so they are distinguished using modules, and every gesture is stored to show on the output screen.

NON-FUNCTIONAL REQUIREMENTS



- **Performance Requirements** - The performance requirements refer to static numerical requirements placed on the interaction between the users and the software.
- **Response Time** - Average response time shall be less than 2 sec.
- Start-Up/Shutdown Time - The system shall be operational within 20 seconds of starting up.
- **Capacity** - The trained model can accommodate 1 hand at a time.
- **Utilization of Resources** - We can add as many gestures we want and train the model, provided we have enough storage capacity or we have to extend the storage capacity.

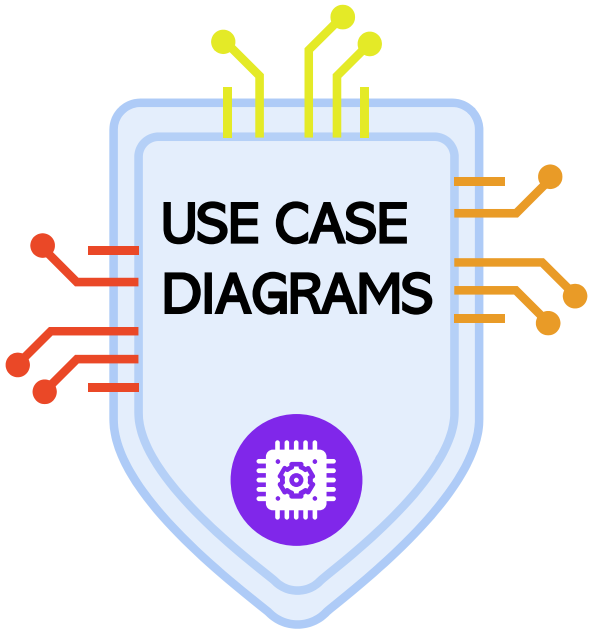
Software Quality Attributes

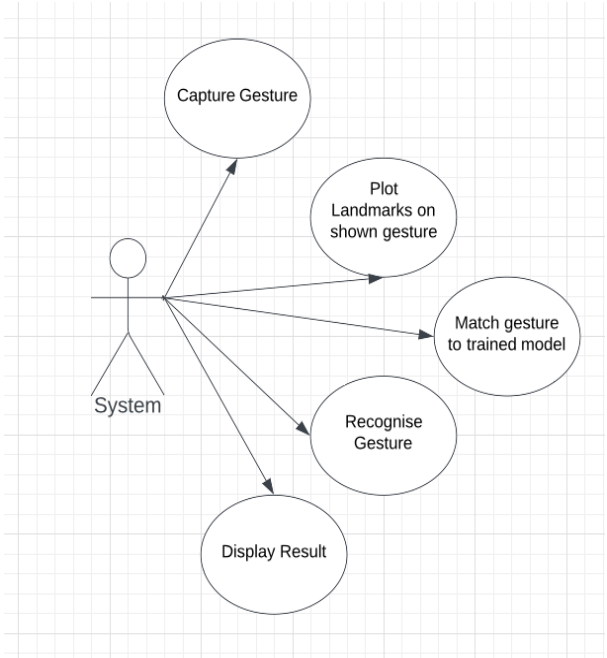
Scalability - The system will be designed in such a way that it will be extendable. We can train and add different gestures as we go.

Availability - The system will be available to all its users round the clock i.e. they can access the system at any time.

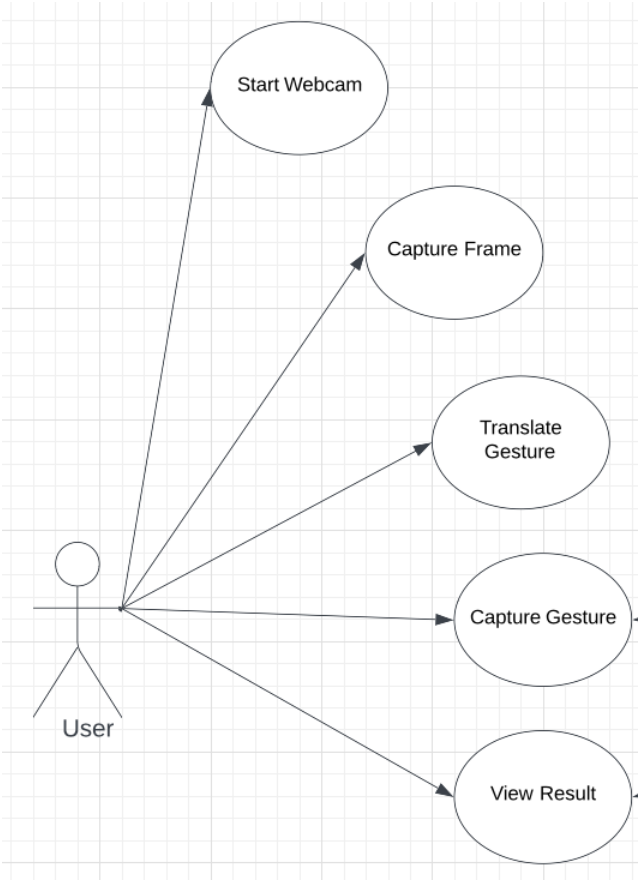
Usability - The interfaces of the system will be user friendly enough that every user will be able use it easily.

SYSTEM DESIGN



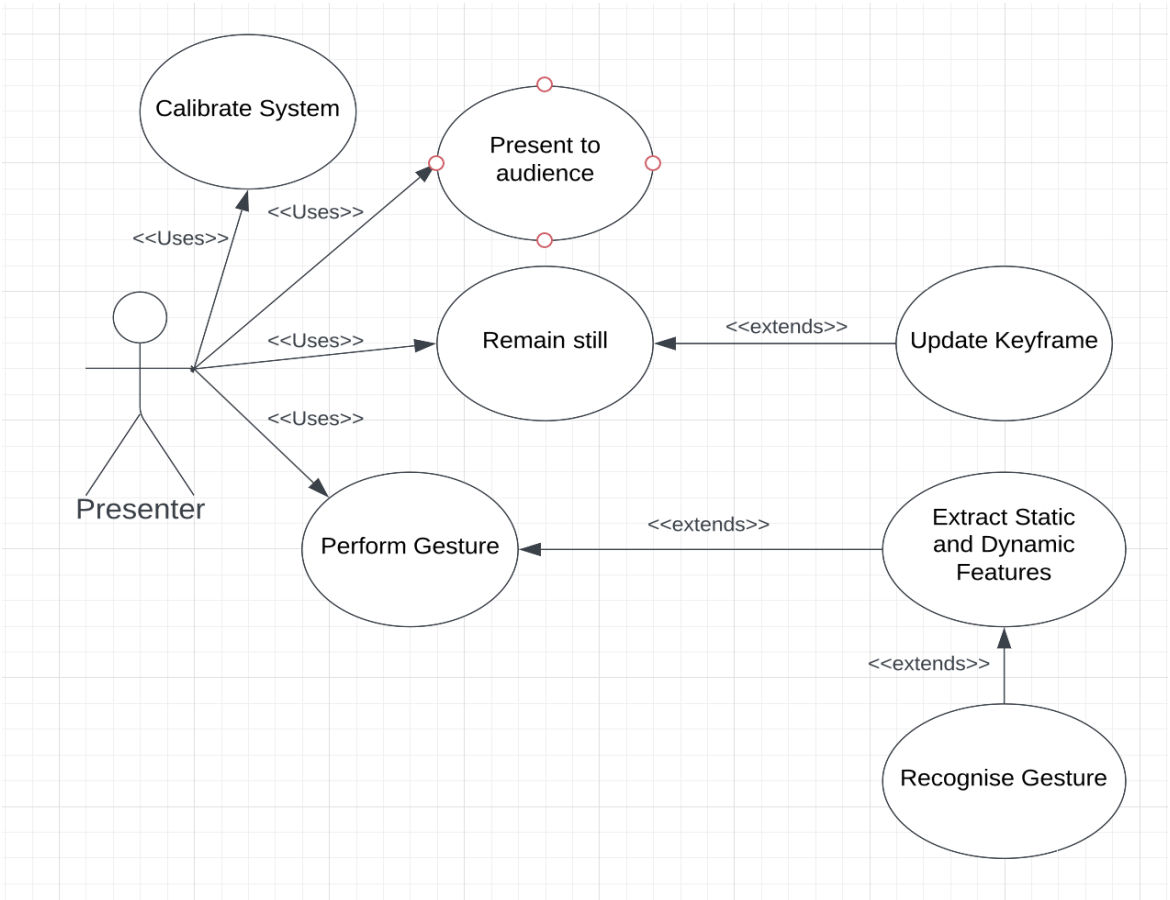


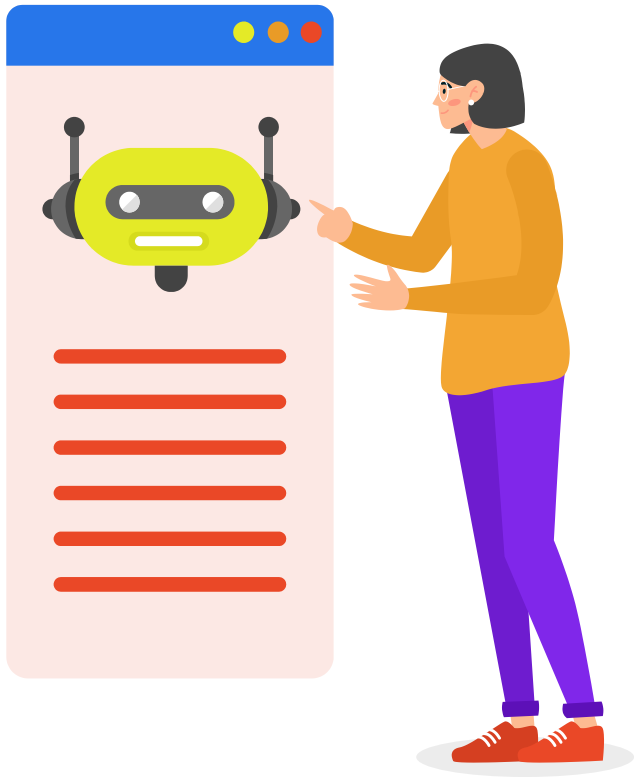
i) USE CASE DIAGRAM FOR SYSTEM



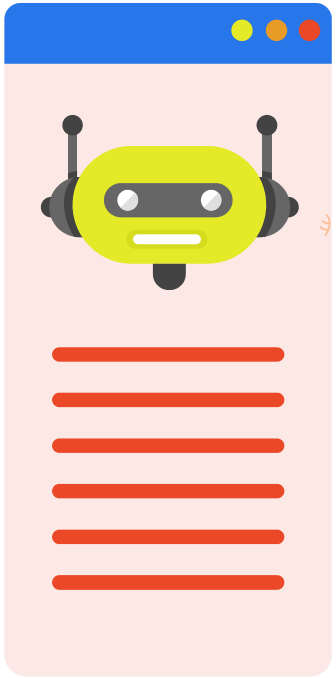
ii) USE CASE DIAGRAM FOR USER

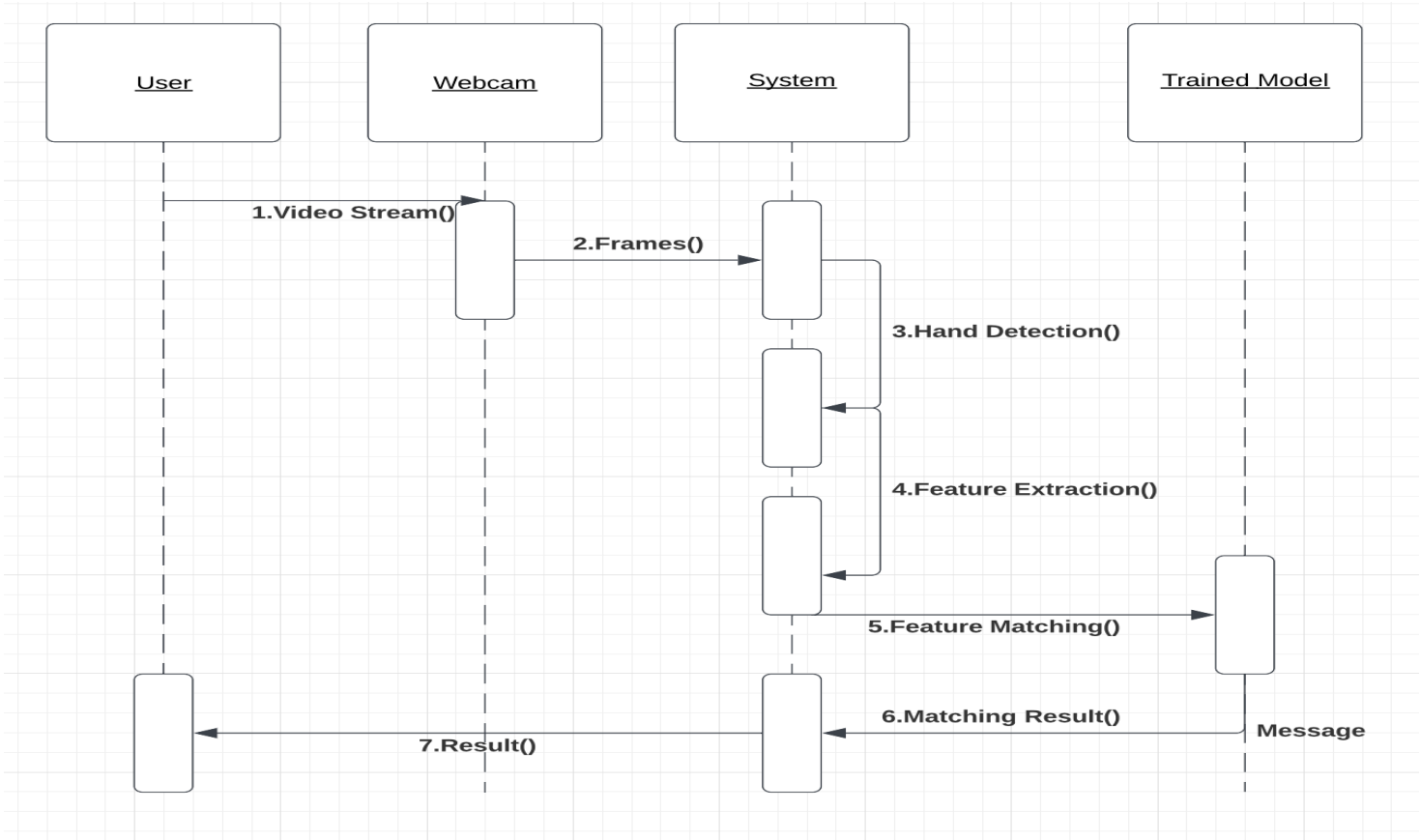
iii) USE CASE DIAGRAM FOR PRESENTER



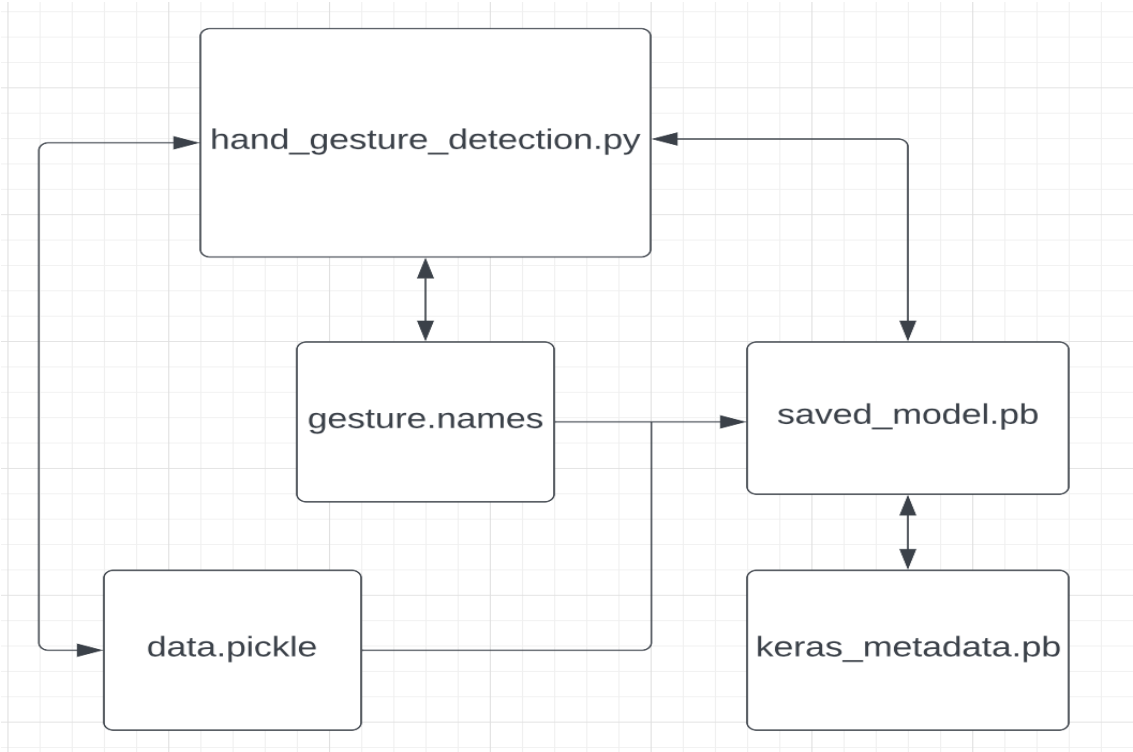


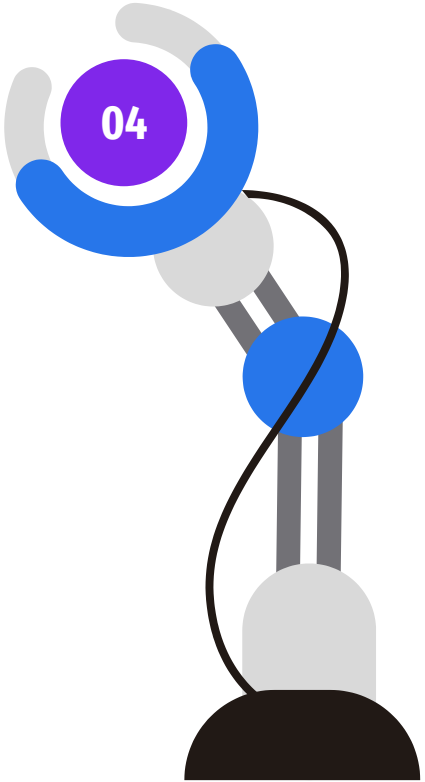
SEQUENCE DIAGRAM



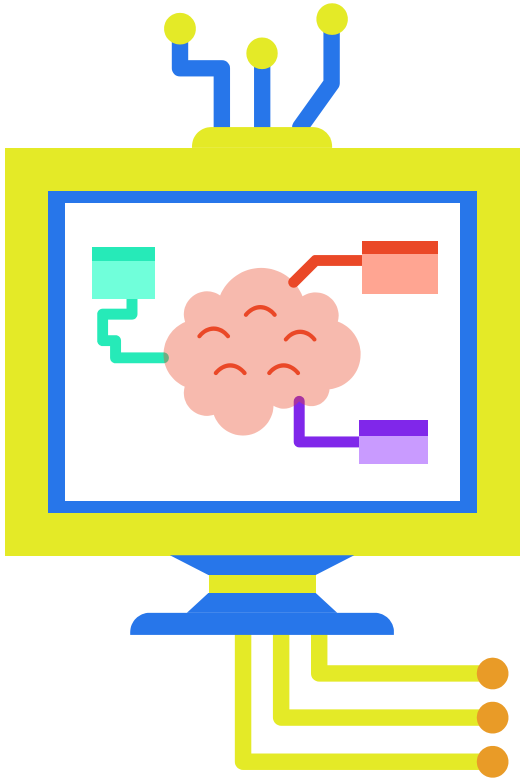


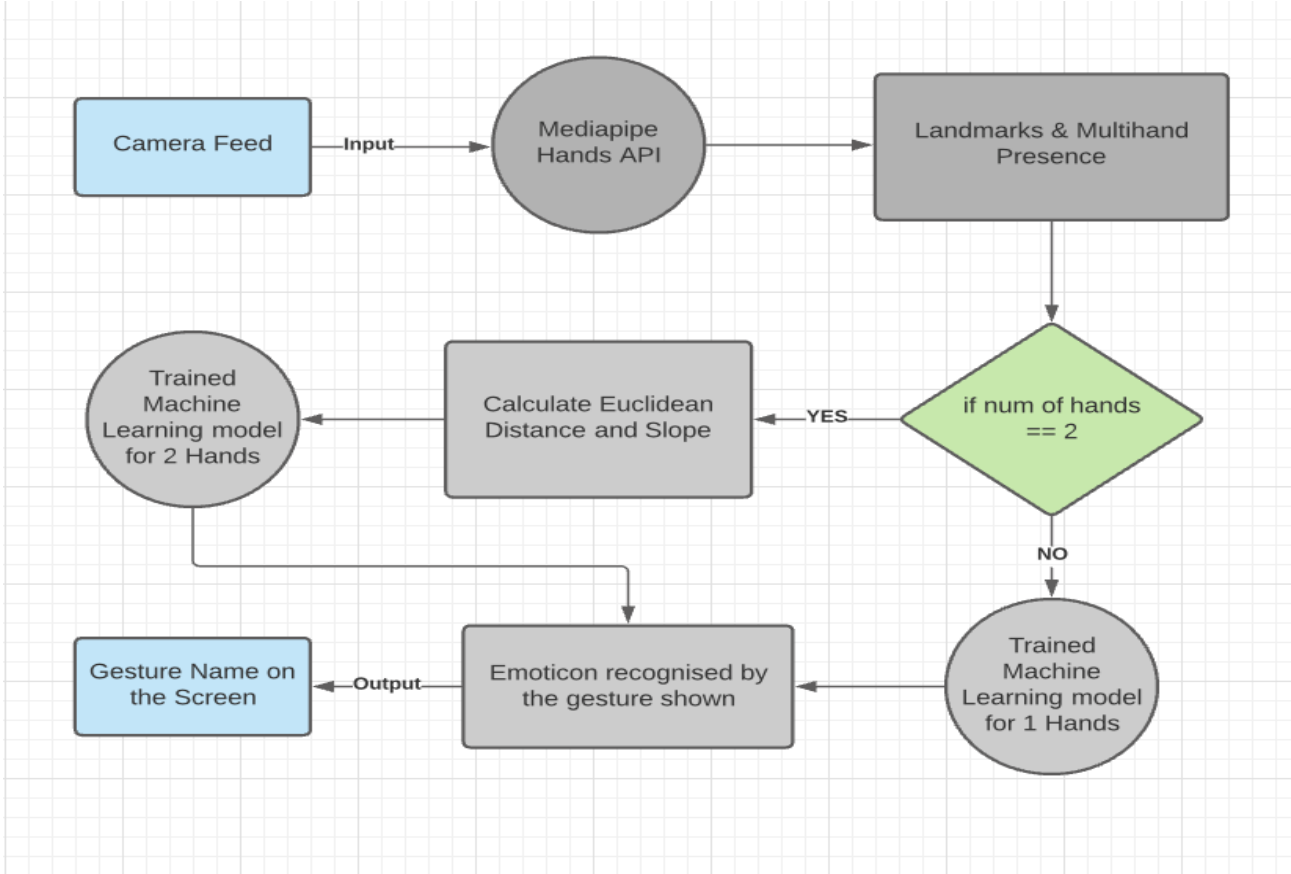
CLASS DIAGRAM



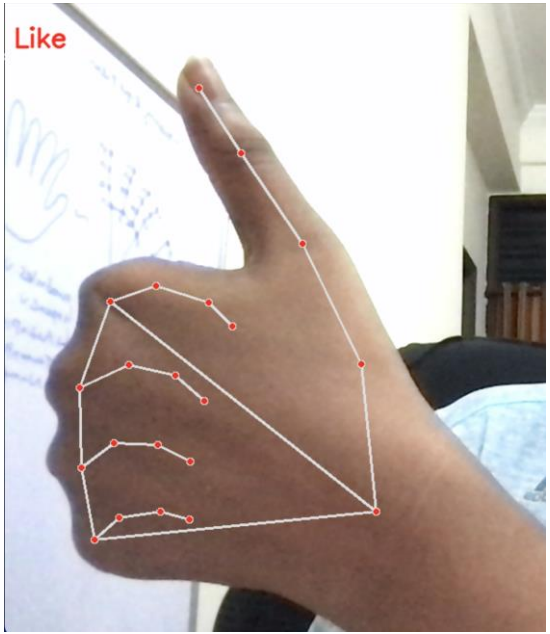


PROCESS FLOW DIAGRAM

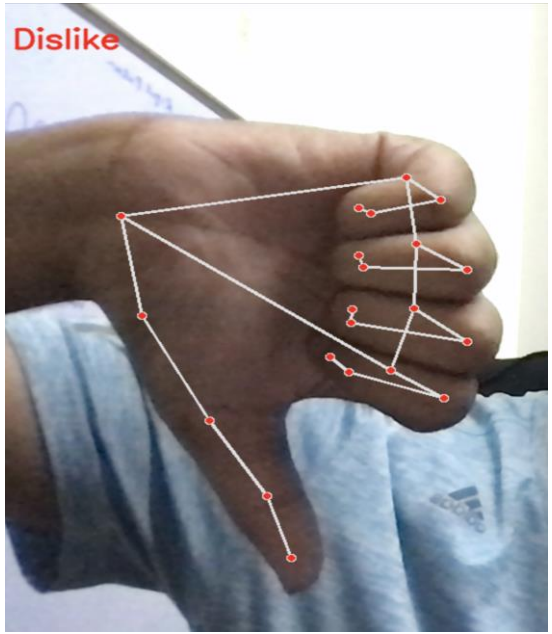




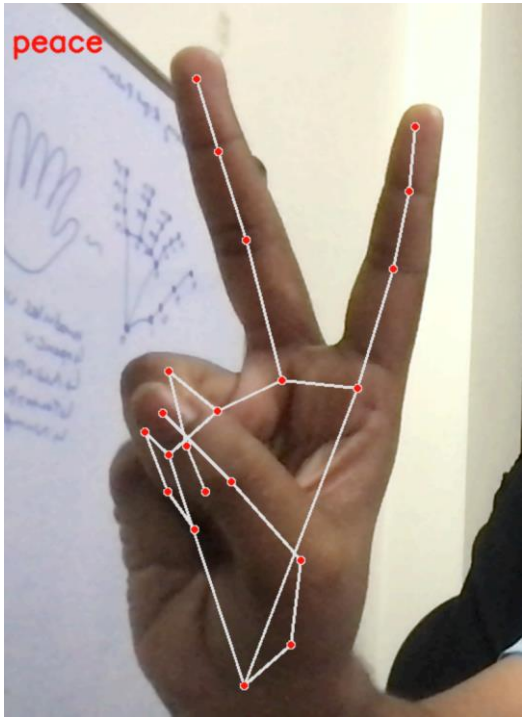
APPLICATION PREVIEW



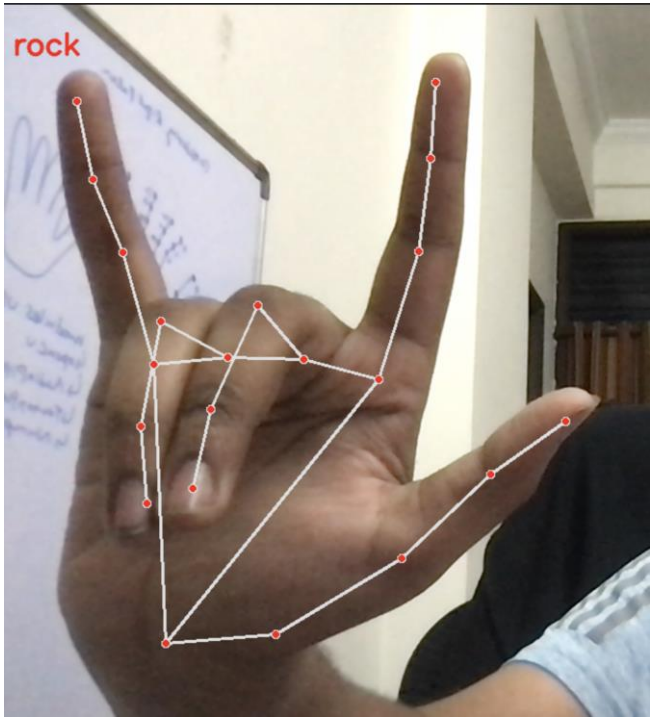
Gesture for Like



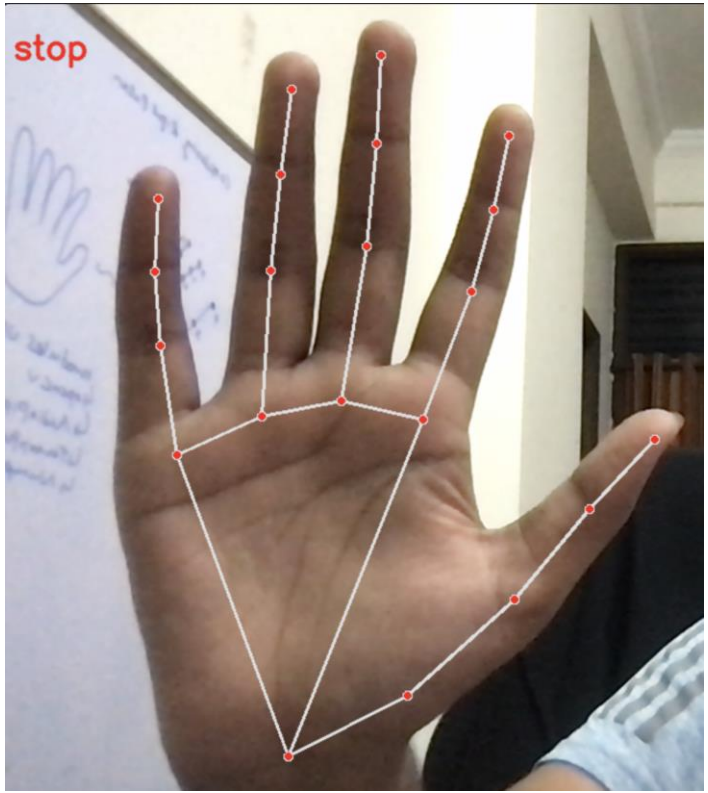
Gesture for Dislike



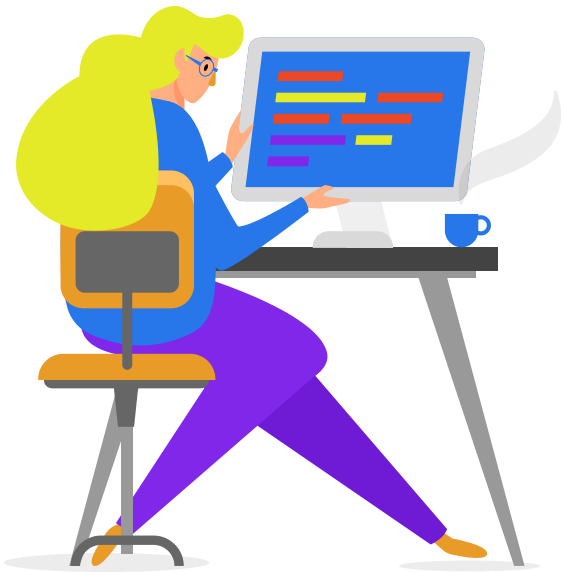
Gesture for Peace



Gesture for Rock



Gesture for Stop



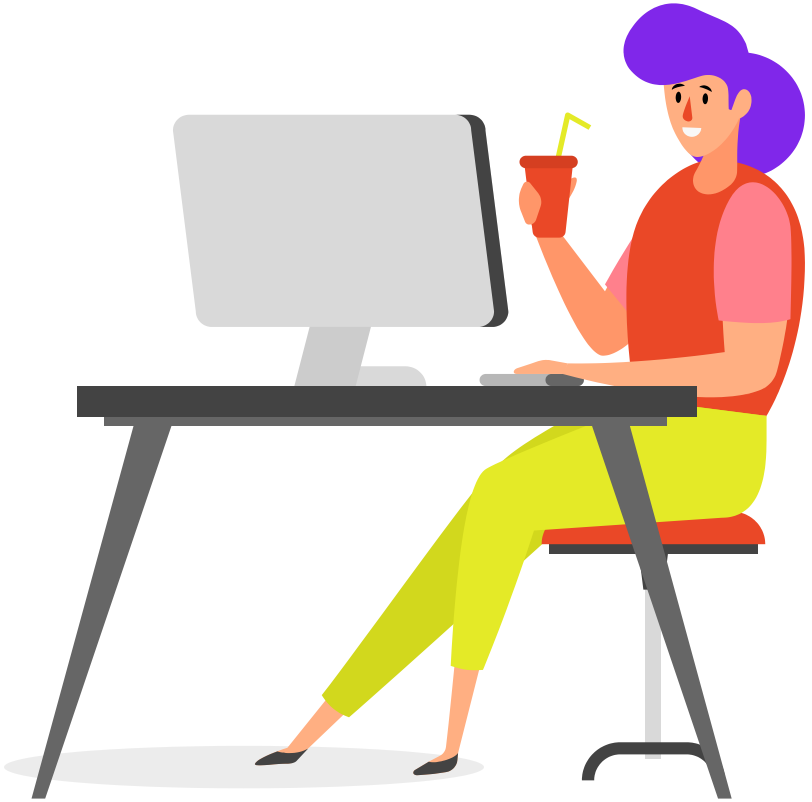
CONCLUSION AND FUTURE SCOPE

- By using computer vision and media pipe we are able to draw landmarks on the hand and by using the KNN model we are able to predict the hand gesture with 95.7% accuracy and fast output.

- The future scope of this project is to break the barrier and act as a medium between the speechless and the people who could communicate. And to also communicate gesture recognition will be able to convey the messages in a hassle free way.

- Our aim is to decrease the response time to predict the gesture in the future iterations of this project and to add an extensive set of gestures for wider usage which can be updated anytime such that our model is up to date.

THANK YOU !



TEAM-04
CSE-B/NGIT