



AI Arena 2.0

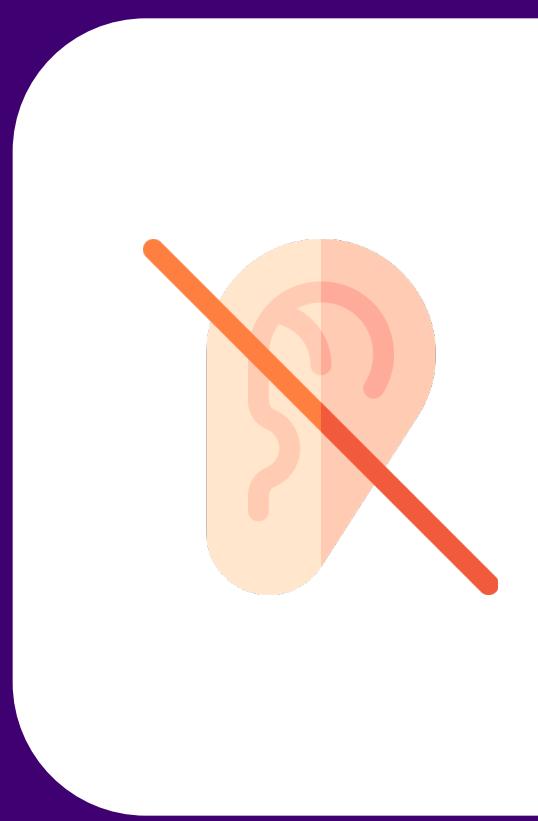
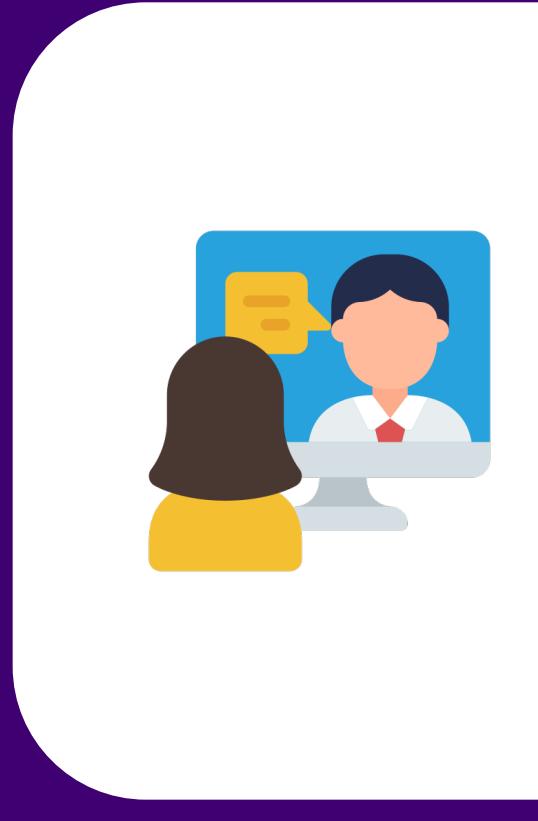
Theme: Social Good



Signify
Real Time Sign Language
Translation App

Aadya Dewangan

PROBLEM STATEMENT



Video calling has become the norm,
especially since Covid

However, Individuals with hearing and
speaking issues face communication barriers.

Present solutions offer real-time captions but
fall short for those who rely on sign language.

BUT, IS IT REALLY A BIG ENOUGH PROBLEM?

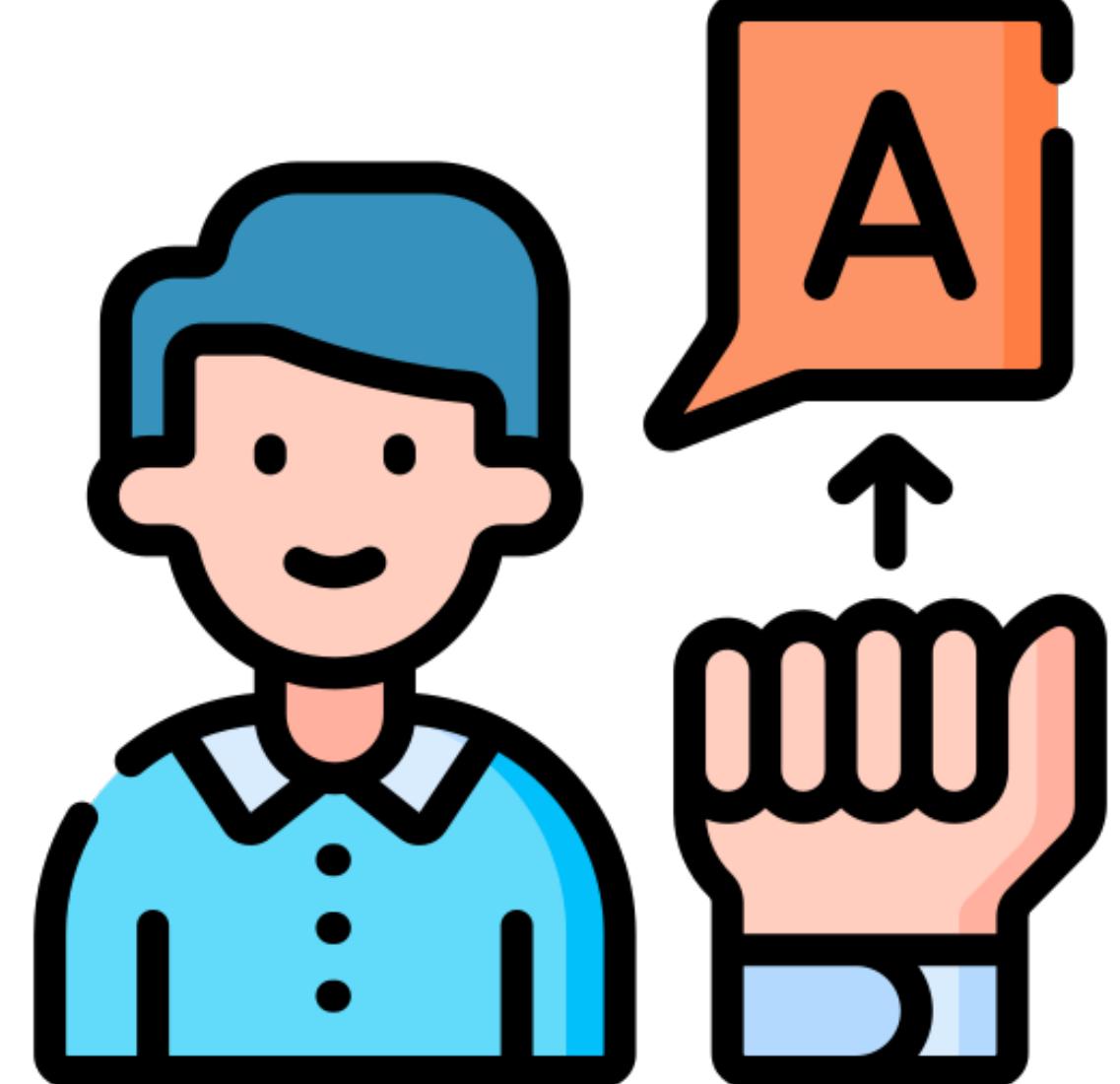
No, it's even bigger than one would think...

According to the World Health Organization, approximately **70 million people in the world are deaf-mutes**. A total of 360 million people are deaf, and 32 million of these individuals are children.

BRIDGING THE GAP: THE SOLUTION

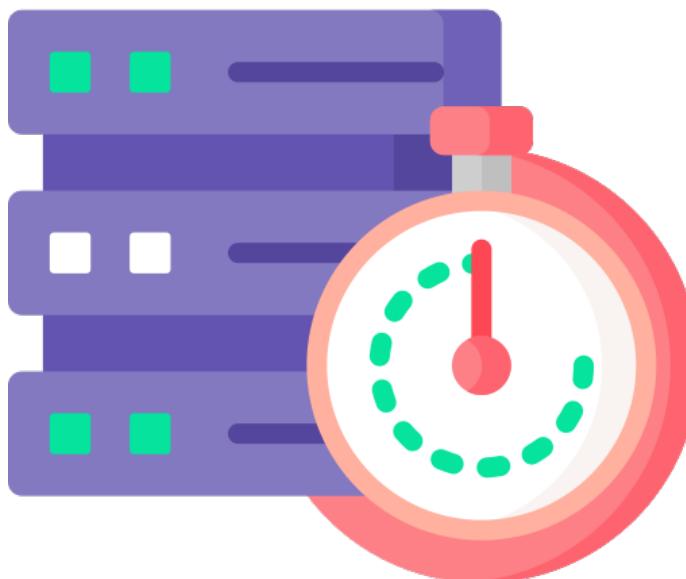
Real Time Sign Language Translation

To empower individuals with hearing and speaking impairments to communicate effortlessly.



KEY FEATURES

Real Time



Like video captions, sign language captions will appear in real time along with so that everyone can contribute equally and seamlessly

KEY FEATURES

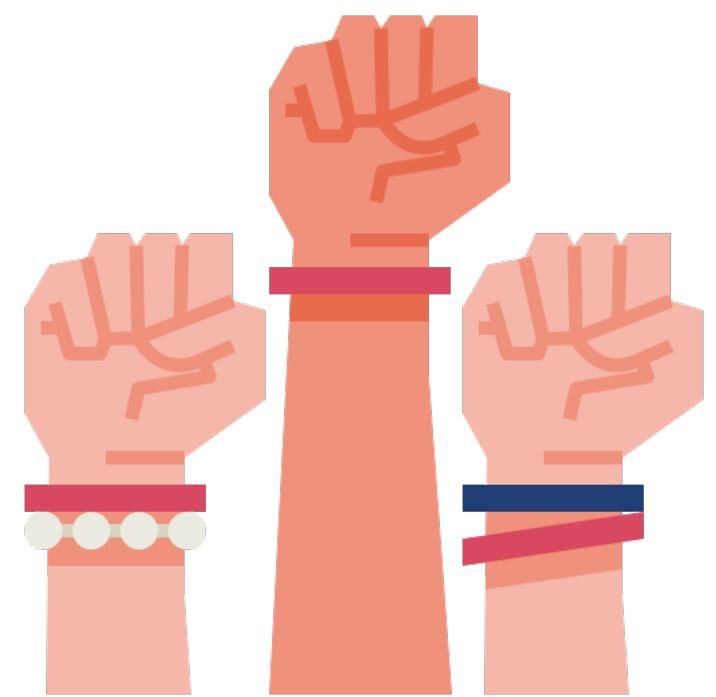
Cross Functional Model



The model is built with Tensorflow and Keras using Google's Mediapipe Library for the detections and predictions. Tensorflow is versatile and robust in terms of deployment, enabling easy integrations.

KEY FEATURES

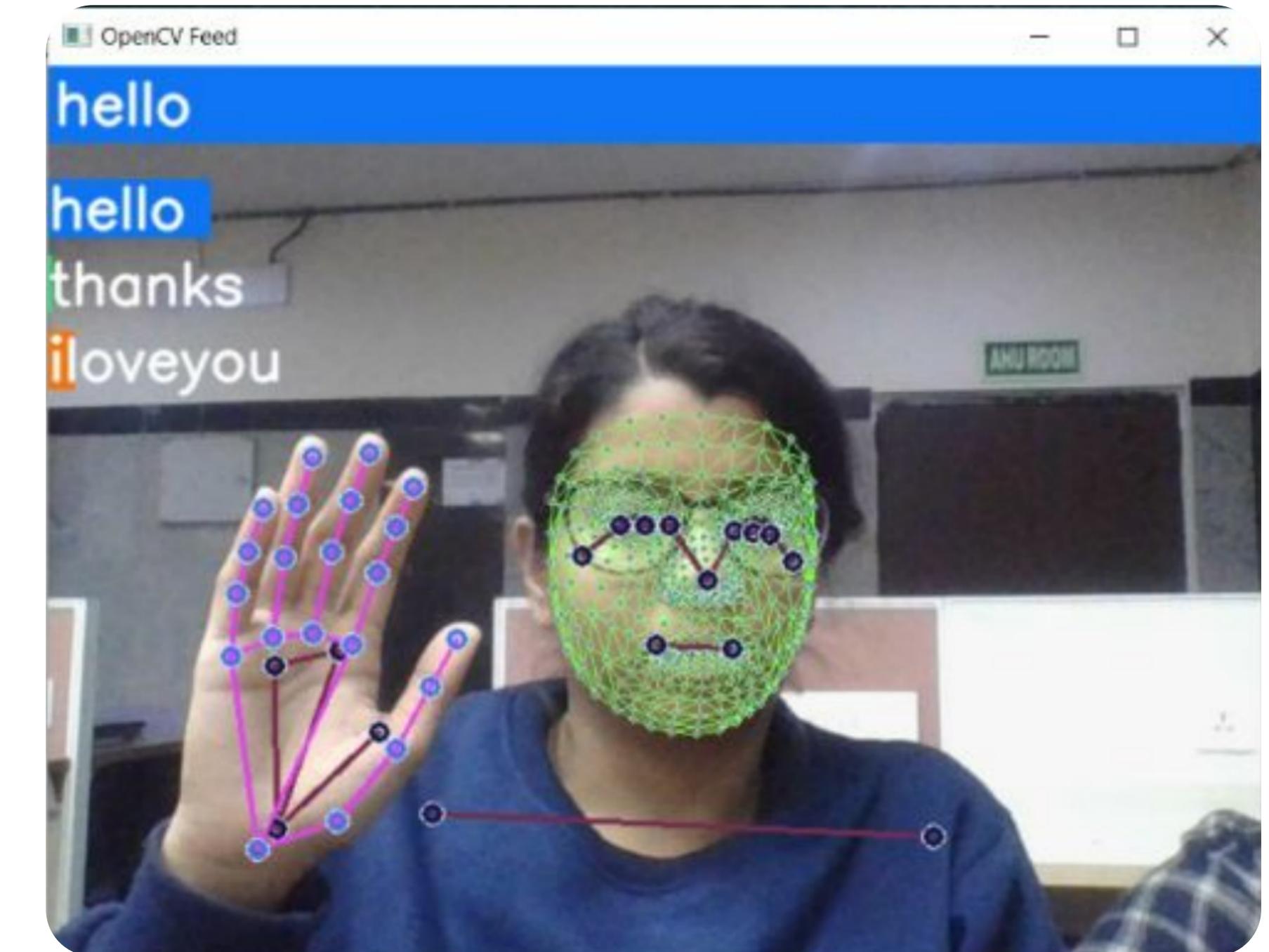
More Empowered Employees



More empowered and happy employees means greater retention rates and greater productivity in the long term and also a better reputation.

PRESENT MODEL

It is presently in Proof of Concept (PoC) stage, with the detection of some basic words.

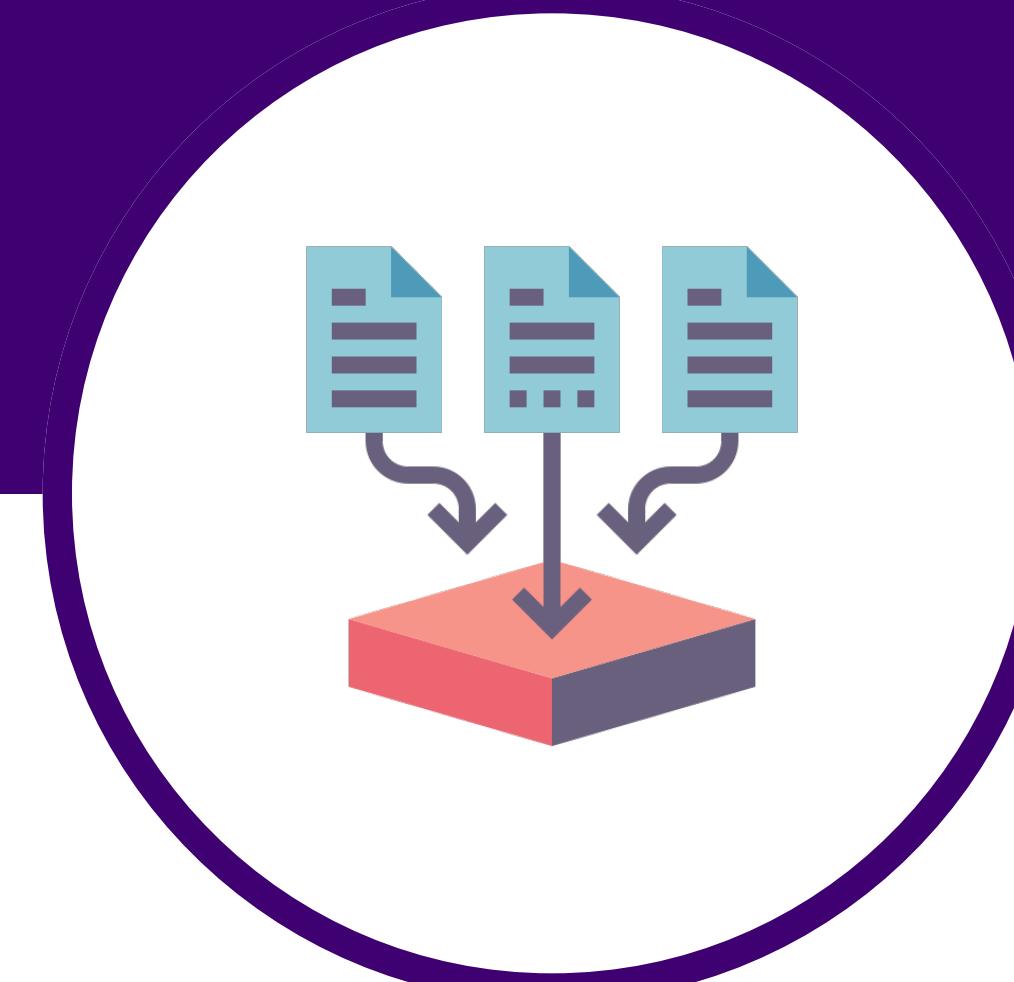


**HOW WAS THE
PROJECT MADE?**

APPROACH TO PROBLEM



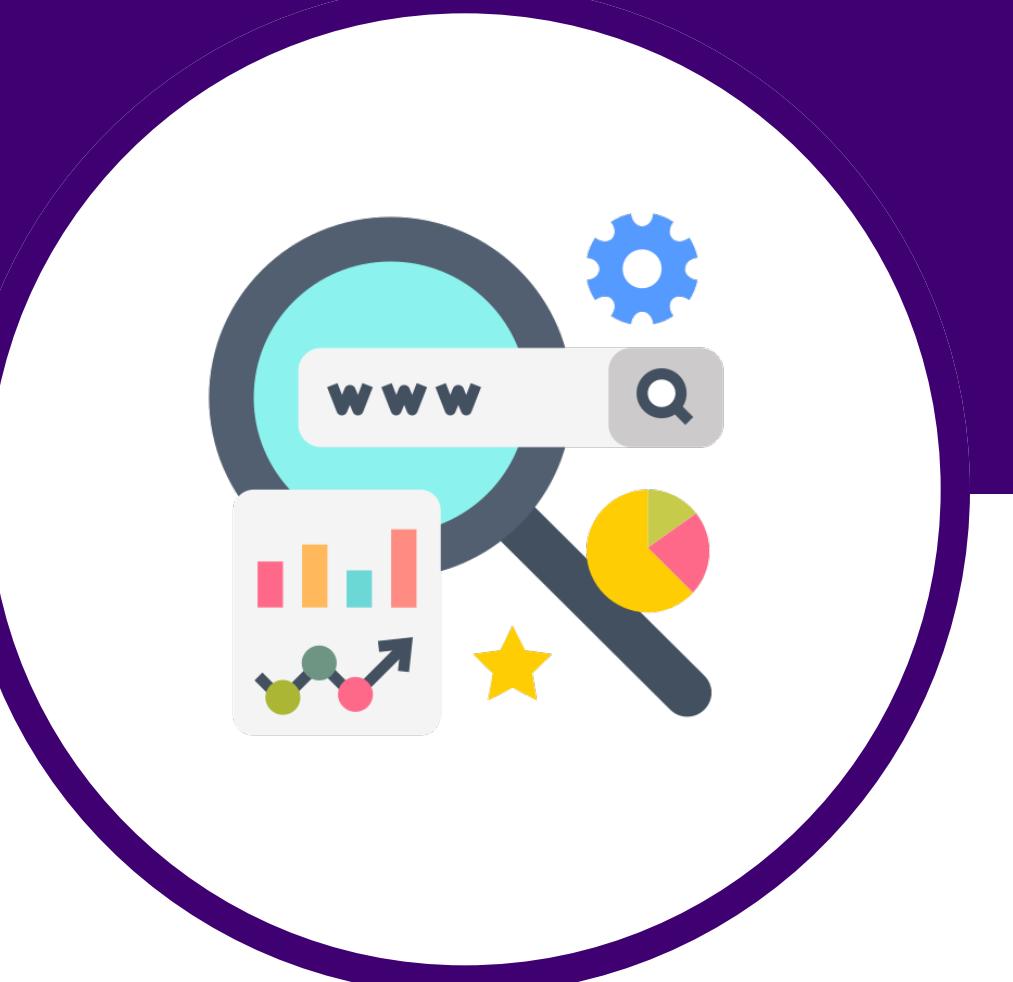
**Detecting
Hands**



**Collecting
Data**



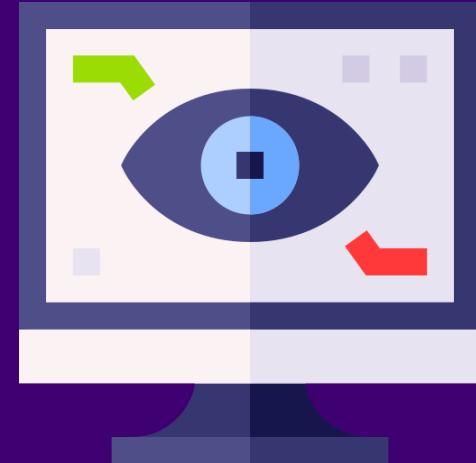
**Model
Training**



**Analysis
of Model**

CONCEPTS USED

**Computer
Vision**



**Long Short Term
Memory (LSTM)**



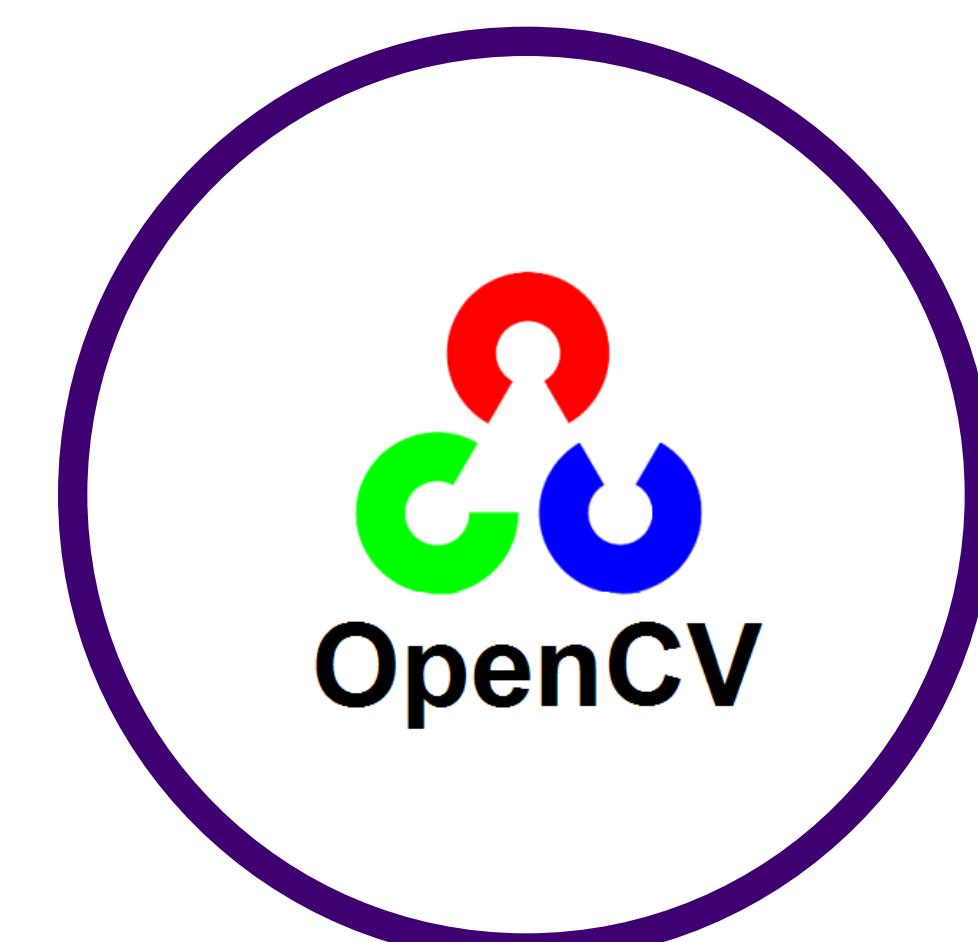
**Data Collection
and Processing**



LIBRARIES USED



to build the model

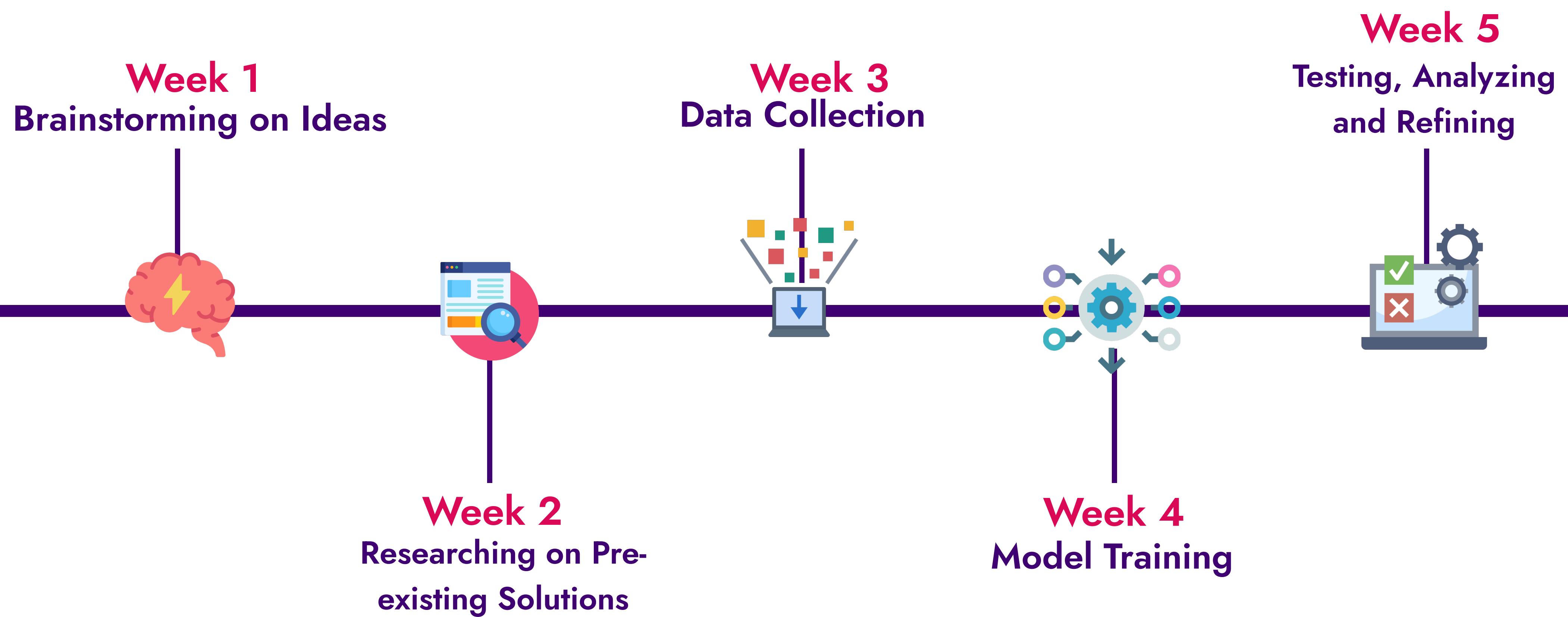


to use the camera to get
information and process it



to detect the gestures

TIMELINE



CHALLENGES

Simple OpenCV model + CNN gives low accuracy and works only on static images

Data not available in needed format

Deployment

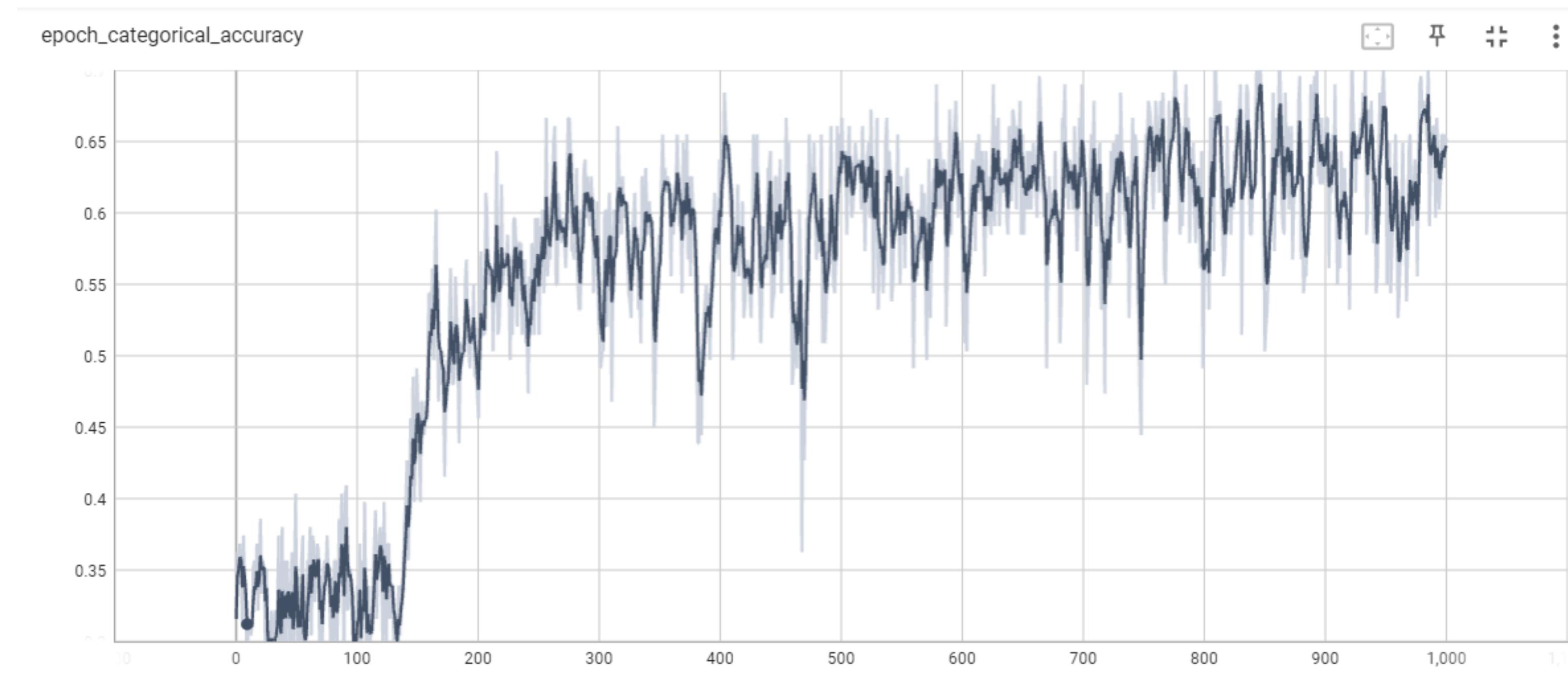
RESOLUTION

Google's MediaPipe Library is specially built for detecting hands, face, pose, etc + LSTM gives a much better accuracy

Created a Data Collection & Preprocessing Flow

I tried creating a Flask app but there were still some issues in the end

MODEL TRAINING GRAPH

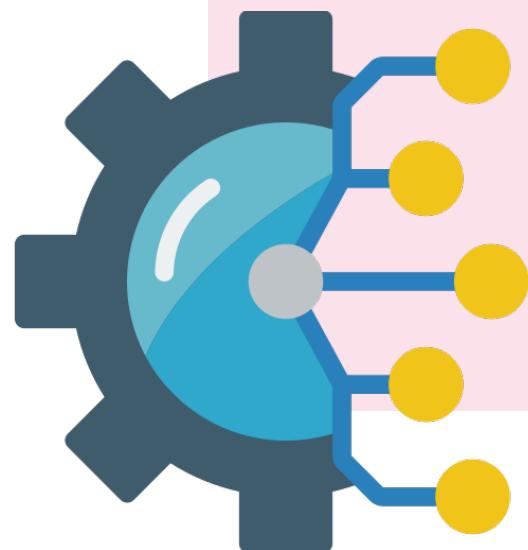


THE FUTURE

POSSIBILITIES...

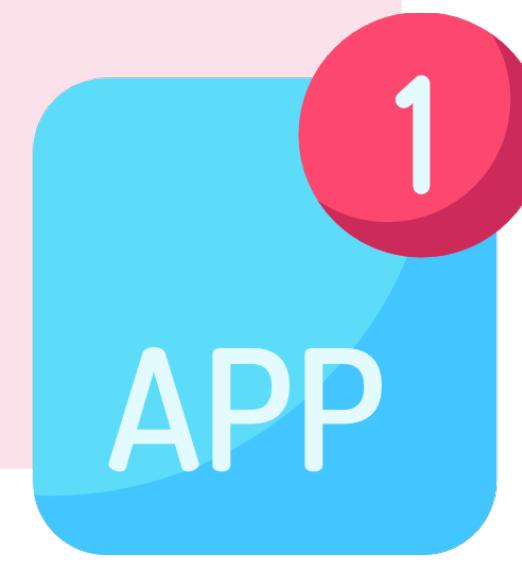
Add on Tool/Extention

The solution can be used as an extension in video calling apps so that these apps can become more accessible.



Stand-alone App

The solution can be used as an app where people can use it in everyday situations without video calling.



CHALLENGES



Data Collection

Available Data is fragmented and would need to be collected properly to train the model.



Sign Language Structure

Sign Language Structure is slightly different than spoken language and might require tweaking of pre-existing LLMS so that they can be integrated together.



Variability Across Languages

Like spoken language, sign language varies across various regions and nations.



Privacy and Security

As videos will be collected, privacy and security must be maintained.

THANK YOU!