Gesture Sync

Presented to: Mr. Pankaj Raj Dawadi

START

Our Team

Prasiddhi Dahal

Kriti Gautam

Samip Pokhrel

Dilasha Upadhyay

Introduction

- For Specially abled people
- A web App that translates sign language to words
- Teach more people about sign language through easy-to-access education on our platform.

Technical Background

1. FRONTEND
   - HTML
   - CSS
   - ml5.js
   - p5.js

2. BACKEND
   - Node.js
   - Express.js

Related Works

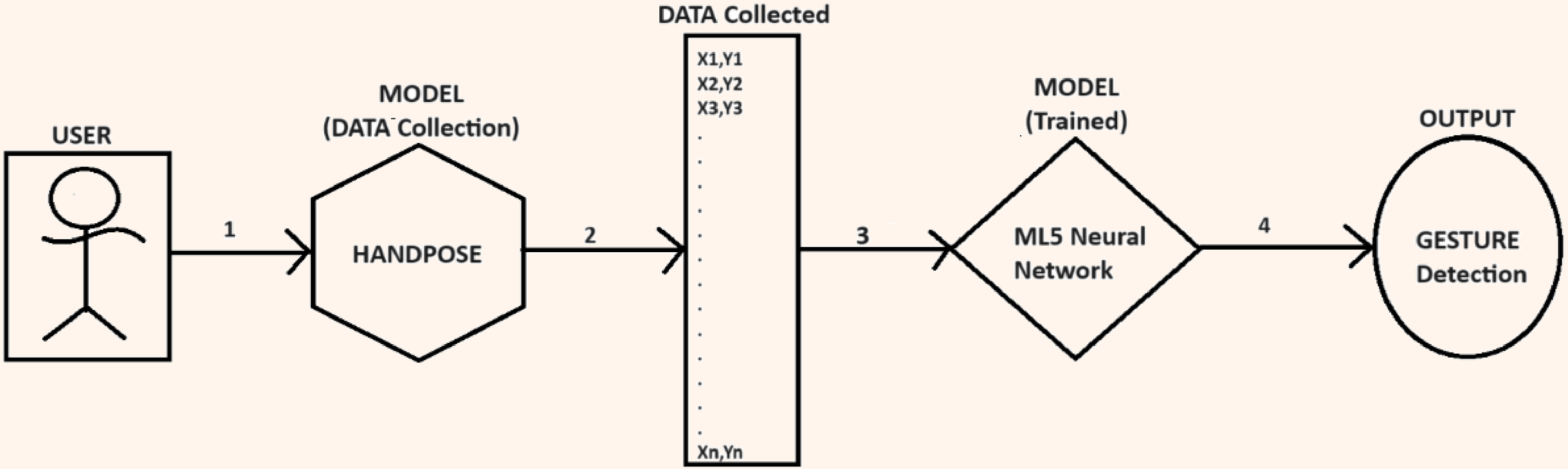
Hand Talk Translator

The Hand Talk app automatically translates text and audio to American Sign Language (ASL) [Beta] and Brazilian Sign Language (Libras) through artificial Intelligence.

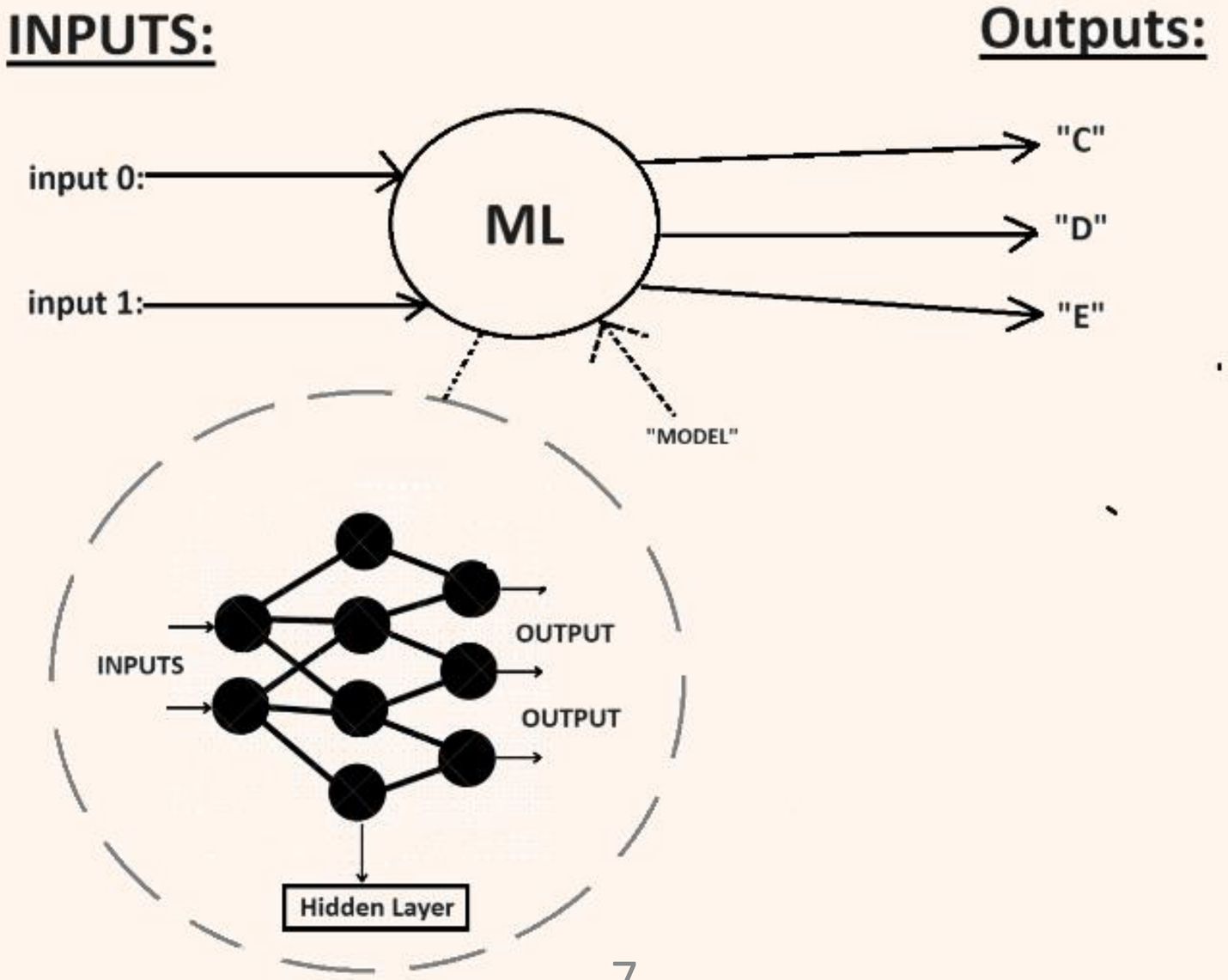
SignAll

The technology, based on computer vision and natural language processing (NLP), and augmented by artificial intelligence (AI), can accurately translate American Sign Language (ASL) into English.

Project Structure



Working of Neural Network



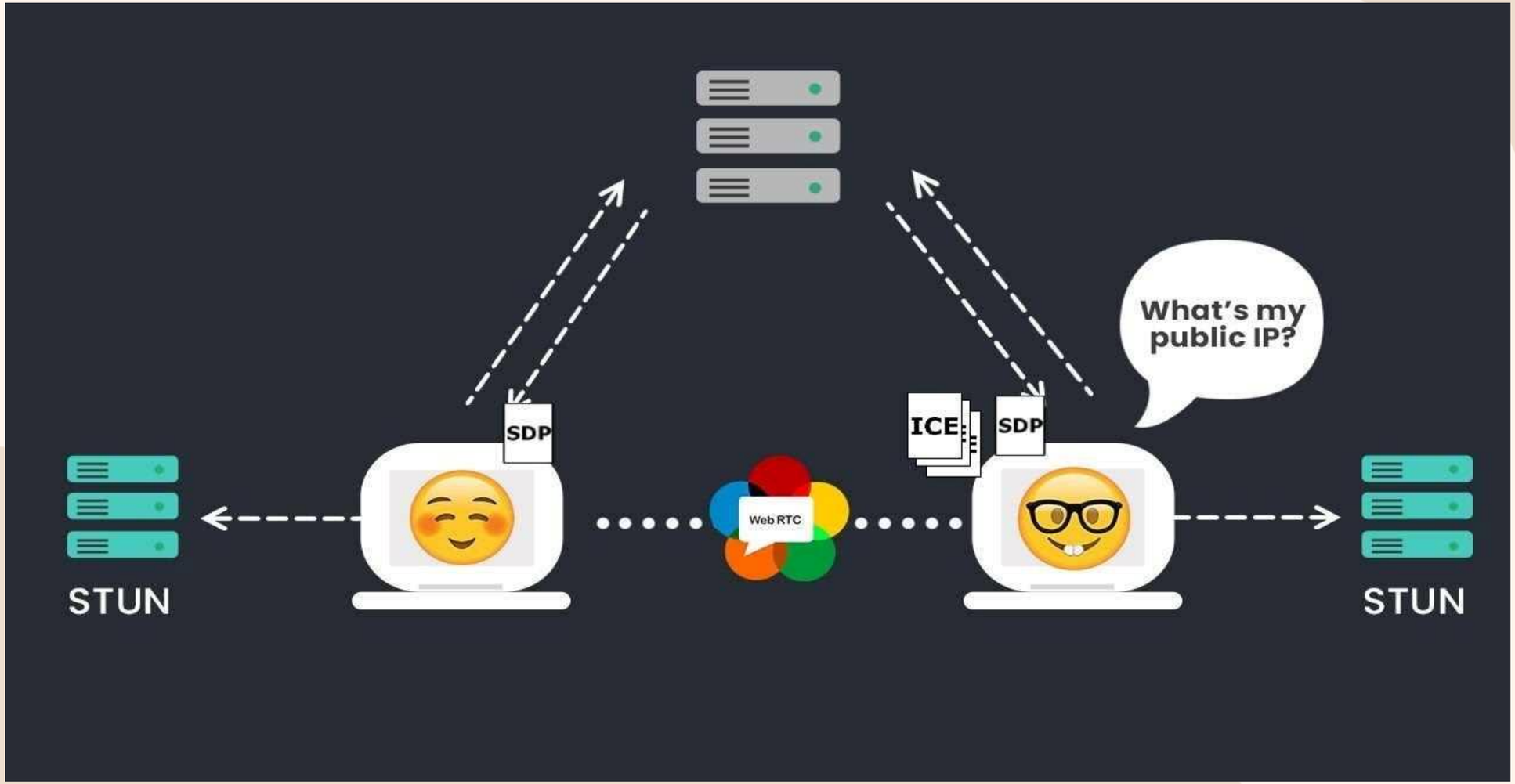
```
annotations:
  indexFinger: Array(4)
    0: (3) [358.6619856149929, 179.24219879429202, -8.030757904052734]
    1: (3) [393.5107517394813, 183.0000268531561, -28.20528793334961]
    2: (3) [357.5866244989198, 219.4375550699324, -27.246686935424805]
    3: (3) [339.0120401293223, 229.5516620471284, -15.61642837524414]
    length: 4
    [[Prototype]]: Array(0)
  middleFinger: (4) [Array(3), Array(3), Array(3), Array(3)]
  palmBase: [Array(3)]
  pinky: (4) [Array(3), Array(3), Array(3), Array(3)]
  ringFinger: (4) [Array(3), Array(3), Array(3), Array(3)]
  thumb: (4) [Array(3), Array(3), Array(3), Array(3)]
  [[Prototype]]: Object
```

```
(10) [{…}, {…}, {…}, {…}, {…}, {…}, {…}, {…}, {…}, {…}]
  0: {A: 0.9999666213989258, label: 'A', confidence: 0.9999666213989258}
  1: {C: 0.00003175268648192865, label: 'C', confidence: 0.00003175268648192865}
  2: {G: 0.00001532396140646597, label: 'G', confidence: 0.00001532396140646597}
  3: {D: 6.808387098544699e-8, label: 'D', confidence: 6.808387098544699e-8}
  4: {H: 2.1147098294704847e-8, label: 'H', confidence: 2.1147098294704847e-8}
  5: {B: 2.3364309618967294e-12, label: 'B', confidence: 2.3364309618967294e-12}
  6: {F: 4.682751582330003e-13, label: 'F', confidence: 4.682751582330003e-13}
  7: {I: 3.6889538461686716e-14, label: 'I', confidence: 3.6889538461686716e-14}
  8: {J: 2.3818735883380376e-16, label: 'J', confidence: 2.3818735883380376e-16}
  9: {E: 4.1635512516222965e-20, label: 'E', confidence: 4.1635512516222965e-20}
  length: 10
  [[Prototype]]: Array(0)
```

```
landmarks: Array(21)
  0: (3) [202.26421035287396, 252.64627571670974, -0.0008528456091880798]
  1: (3) [270.1171503814829, 272.5623713333167, -8.359554290771484]
  2: (3) [343.6990250873173, 262.9923599649169, -15.813541412353516]
  3: (3) [395.4736820513865, 245.91786703891395, -24.574920654296875]
  4: (3) [425.555869126228, 236.03374138846172, -32.98649978637695]
  5: (3) [358.6619856149929, 179.24219879429202, -8.030757904052734]
  6: (3) [393.5107517394813, 183.0000268531561, -28.20528793334961]
  7: (3) [357.5866244989198, 219.4375550699324, -27.246686935424805]
  8: (3) [339.0120401293223, 229.5516620471284, -15.61642837524414]
  9: (3) [327.1473219104037, 151.2906666180569, -12.114435195922852]
  10: (3) [359.91084255318685, 160.15895195301783, -35.62973403930664]
  11: (3) [315.59948977502404, 209.63745388197262, -34.388397216796875]
  12: (3) [297.93923106749423, 212.46219061558799, -23.307422637939453]
  13: (3) [288.4897492556038, 129.85846809858796, -18.775880813598633]
  14: (3) [321.15212899595105, 136.3194366597057, -40.89241027832031]
  15: (3) [282.67384843961906, 186.61552715849962, -38.56059646606445]
  16: (3) [265.09477877206956, 191.0321072230386, -29.33123016357422]
  17: (3) [245.6778892313195, 115.09947306723566, -25.583179473876953]
  18: (3) [274.49254915898575, 115.39929245426684, -41.30586242675781]
  19: (3) [257.36408060037417, 156.20723339198315, -37.14340591430664]
  20: (3) [244.36740422547095, 166.76669760081717, -26.73927879333496]
  length: 21
  [[Prototype]]: Array(0)
```

Gesture Recognition Video Chat Implementation

(USING webRTC)



OVERALL

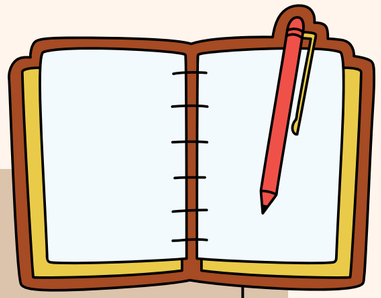
FEATURES

- Sign Language Detection
- Train and Download own model
- User-Friendly UI

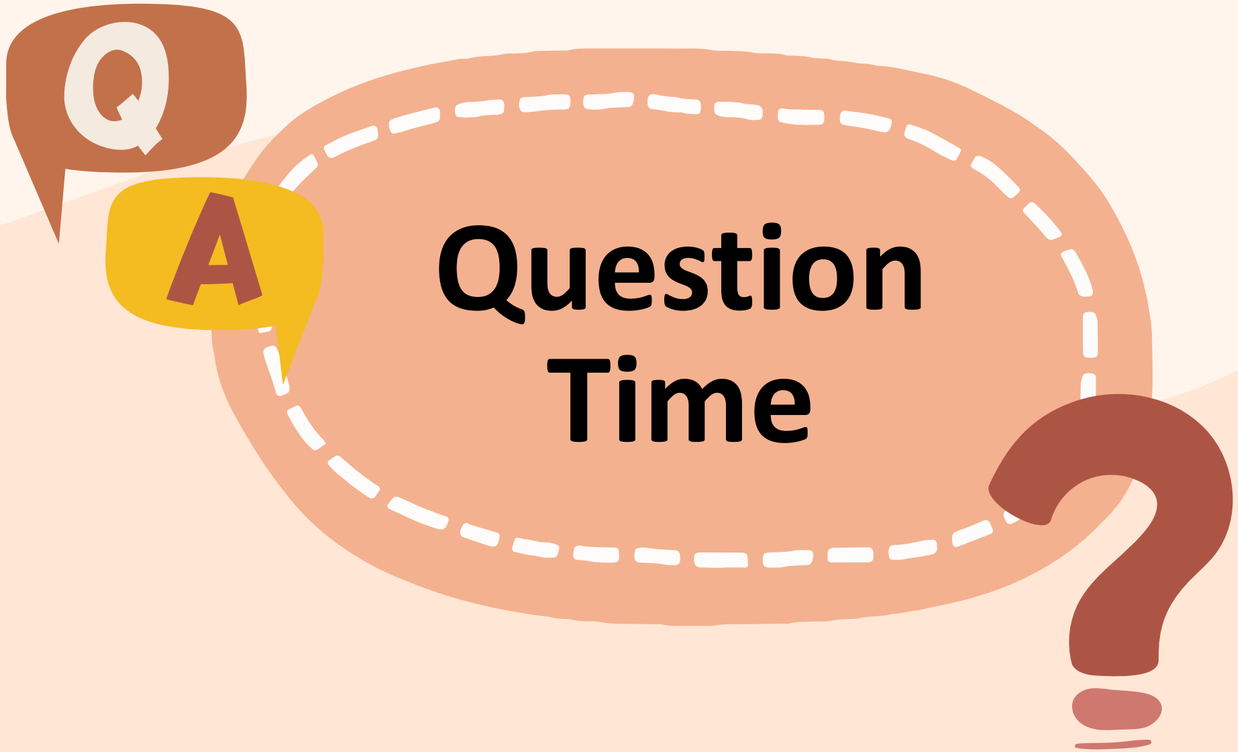
CHALLENGES

- Integrating ML model in website
- Constrained size of our trained dataset

Conclusion



Our project is ML based gesture detection system in website, focused on solving the language barrier problem between people with and without voice. While we are still working to make it more flexible and accessible for all, we learned a lot during the whole process of searching and trying different solutions.



Question Time

THANK YOU SO MUCH!