Names: Oren Raz and Lam Kai YI Karen

Assignment: REI603M Assignment 5 – Product Development for Sign to Text model

**Description of typical User – User with hearing impairment**

Jón is a 25 years old male with hearing impairment. He has an outgoing personality and enjoys the company of friends. As a child, Jón studied in a school for the hearing impaired, where people around him comminucated in sign language. He enjoyed the school environment as he could freely and independently communicate with his friends and teachers.

However, since his graduation from school, Jón realizes that he is having difficulty interacting with the wider community, as they are unable to understand sign language. He misses the time in school, where he was surrounded by people with the knowledge of lip reading and sign language.

Trying to adapt to his new environment, Jón uses his ability to lip read to understand what his co-workers are trying to convey. However, not all his co-workers are well-versed with understanding sign language. Hence, expressing his thoughts to others is challenging. Jón would either use sign language to communicate or must resort to typing is message on his phone. This progress is very slow and limits conversations to short and shallow ones.

When Jón heard about our project initiative, he was very excited by what the application could offer. The application could provide him with a way to communicate with anyone he meets. All he needs is to whip out his phone, open the app, place it in front of him and speak! The application provides an accurate translation of his signed words to text for seamless communication between him and his co-workers. The ability to enage in a conversation naturally without having to use type out his thoughts makes him feel more included in the community.

Jón understands that it is still in development and is far from perfection. However, he sees the potential the application has in helping people like him. He hopes to provide feedback to improve the application and patiently looks forwards to each version update. One area he hopes the application can improve on, would be to reduce the lag time between his signed words and the translated text appearing on the screen, yet there is a more important improvement he hopes for, that is, significantly reducing the amount of incorrectly translated words.

**Step By Step Guide on User Interaction with Product**

The product will be a mobile application with a simple and easy-to-use user interface. The product will feature a camera function, which captures real-time video of the user. When the user signs a word, the video is immediately recorded and sent to cloud to be put through a pre-trained machine learning model, instantaneously converting the signed word into text. This text appears on the phone screen for the non-signer to understand. The product hopes to bring greater convenience to the hearing impaired and mut community when communicating with the general population. However, understandably, due to the imperfections of the machine learning model, there could be accuracy limitations and lag time in translations. These are areas which the application development team hopes to continuously improve.
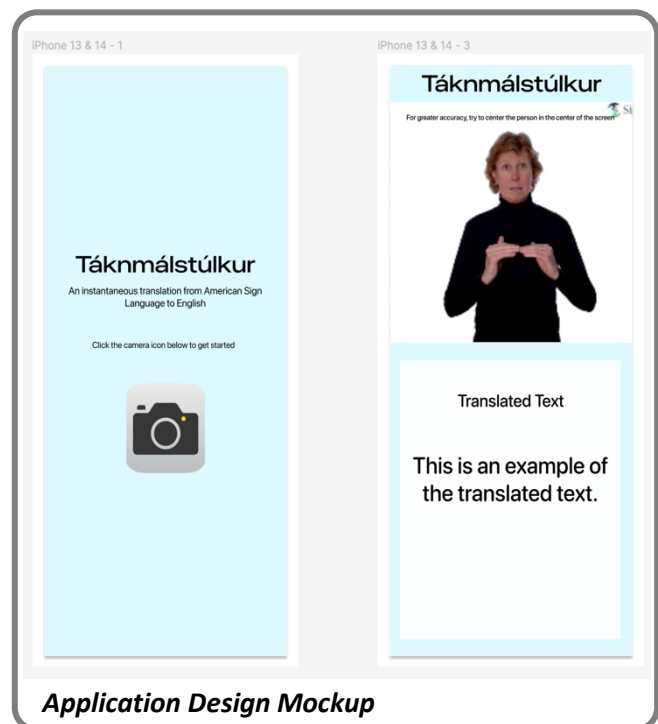
The product will also be available as a PC application, for any modern PC with a GPU and a webcam, designed to work without internet connection. Public offices will be interested in installing it in any service reception office, where hearing impaired people may get services related to their private information, with a smaller risk of information leakage.

**Our Current Progress:**

- Data collection of 11,980 videos labelled by 2000 words.
- Data exploration and cleaning – removed missing videos, checked video quality, and re-assigned train, validation and test sets.
- Embedded Data – used mediapipe library to extract key landmark points (543 landmark points per frame in video, each with x, y, z coordinate).
- Data Engineering – Creation of more train data by moving individual landmark points in data and laterally flipping landmark points.

**In Progress – Model Engineering:**

- Attempted using mean value across all frames, followed by Logistic Regression, Support Vector Machine and Random Forest.
- Attempted Dynamic Time Wrap (DTW) followed by k-means.



*Application Design Mockup*

**Weekly Plan**

| 26 Feb (Mon) – 1 Mar (Fri) | Model Engineering: <ul><li>Attempt Recurrent Neural Network – LSTM on subset of train data set.</li><li>Further experiment with DTW, followed by other classification models.</li></ul> |
|---|---|
| 4 Mar (Mon) – 8 Mar (Fri) | Model Engineering: <ul><li>Compare the validation results of models (with larger training data size) - including Data Engineering.</li></ul> |
| 11 Mar (Mon) – 15 Mar (Fri) | Model Engineering: <ul><li>Conduct Hyperparameter Fine Tuning, explore additional NN structures.</li></ul> UI/UX of Application: <ul><li>Work on user interface of application.</li></ul> |
| 18 Mar (Mon) – 22 Mar (Fri) | Model Deployment <ul><li>Attempt to deploy model on the cloud and use the application to connect it with API.</li><li>Conduct Monitoring and Testing.</li></ul> |
| 25 Mar (Mon) – 5 Apr (Fri) | Easter Break |
| 8 Apr (Mon) – 12 Apr (Fri) | Model Deployment <ul><li>Adjustment to model based on testing results</li></ul> |
| 15 Apr (Mon) – 19 Apr (Fri) | Buffer Week |