CS3237 GROUP 12 Preliminary Project Proposal

By: Lim Hao Xiang Sean, Liu Wei Jie Nicholas, Permas Teo Pek Neng, Glen Wong Shu Ze

Problem Statement
In the midst of the global pandemic, many lectures in NUS are held online via video conferencing tools. Communicating via such software tools alone will never be able to sufficiently replace the experience of a face-to-face lesson, hindering learning. Problems:

1. Inability to run interactive activities to engage students
2. Unnecessary cognitive load in navigating video conferencing interface distracts students from lesson content.
3. Harder to notice and observe visual or non-verbal cues among participants. For example, teachers may find it hard to judge if the lesson pace is too fast/slow as opposed to face-to-face lessons where they can observe students more closely

Solution
We would be using IoT sensors and neural networks to support gesture detection on video conferencing platforms. We believe that the idea of conveying thoughts and emotions of students through physical actions will address the issues listed above. We will confine the scope of our project only to head moving actions.

Gesture detection will help increase the interactivity of lessons held on video conferencing platforms. A simple example would be that students can physically participate in class activities such as answering “yes-or-no” questions by nodding or shaking their heads.

The above example also shows how gesture detection makes it easier for students to communicate their thoughts into action, reducing cognitive load. Instead of *tediously* navigating to the video conferencing platform interface, click ‘chat’, then click ‘yes’ button to imply intent, they can simply perform gestures intuitively.

The gestures from the students will then be displayed to the teachers. Hence, teachers will also be able to get a better sense of the response of the classroom via non-verbal cues picked up from the gesture detection feature to better adapt to different class dynamics similar to a face-to-face lesson.

Specifications

<u>Sensors</u>
To gather data for this project, we will be using the altimeter and the 9-axis motion sensor.

The BMP280 altimeter on our IoT device is an absolute barometric pressure sensor that measures changes in pressure at different altitudes. Using this sensor, we can detect head movements that result in a change of the altimeter reading such as a nod.

The MPU-9250 9-axis motion sensor comprises an accelerometer, gyroscope and magnetometer. Both accelerometer and gyroscope data tell us how the user’s head is moving and rotating, allowing us to track the motion in all directions. Magnetometer gives us the absolute spatial orientation and motion vectors that are perfect for head movement tracking.

The data from BMP280 and MPU-9250 are used as inputs in the model to classify the gesture.

<u>Algorithm(s)</u>
To train our model for this project, we will be using Recurrent Neural Network (RNN) for gesture recognition. More specifically, we will be using Long Short-Term Memory (LSTM) neural networks as our

deep learning method. LSTMs are great at pattern recognition and we feel that gesture recognition is mostly about pattern recognition, hence we decided to go with LSTMs. We plan to train our model individually using our laptop to emulate the cloud before combining our data together.

Cloud

Our 4 TI SensorTag units will be sending the collected data to an AWS EC2 Instance where our deep learning model will be hosted on and trained using collected data. The network protocol used here is MQTT.

Product

The final product that we hope to achieve is a trained model that is able to identify mainly four different gestures:

1. Yes - Nodding of head
2. No - Shaking of head
3. Thinking - Head movement upwards, chin moving up
4. Unsure - Tilting head to either side

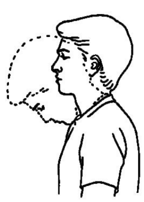
Yes	No	Thinking	Unsure
 Nodding of head	 Shaking of head	 Head movement upwards, chin moving up	 Tilting head to either side

Fig. 1. Gestures

Feasibility

We feel that our project idea is highly feasible since we have restricted the scope of actions to only 4 head gestures and the sensor tag can be attached to the user's head by means of a headband. We can generate the required data for training our model independently by performing variations of those required head gestures and labelling them properly. The TI Sensortag also has appropriate and effective sensors to detect such gestures accurately i.e. 9-axis motion sensor and altimeter. We are also confident of implementing the LSTM model as we have learnt similar variations from lectures.

Roles & Responsibilities

All members will be involved in sensor data collection and model training. Each member will be in charge of and guide the team for the component that he is assigned.

Name	Role	Responsibility
Liu Wei Jie Nicholas	Backend Engineer	Set up communications from gateway to cloud
Lim Hao Xiang Sean	Machine Learning Engineer	Neural network model training
Permas Teo Pek Neng	Data Engineer	Neural network data pre-processing
Glen Wong Shu Ze	IoT Engineer	Setting up of IoT device with cloud