RESEARCH PROJECT PROGRESS

Week 06

Problem Statement

So far, the mainly problem is reading raw data and data manipulation, since the data set I am using is the EEG data. The EEG data is generated from the complex EEG recording system. The following main problems are listed:

1. The original dataset includes 17 patient's experiment's data, and each experiment includes two trials. This dataset would be quite large.
2. The dataset structure is presented by using multiple files, such as files to record channel info, files to record EEG voltage data, and files to record coordsystem.
3. Files formats are different between different EEG machine. It could be a challenge if any vender using a closed system.

Source data

- A electrogencephalogram (EEG) scan data from one participator while playing an 8-bit style video game

- In one data set of one participator, there are two experiment recording (1st and 2nd). In each recording, there are several related file: *.tsv, *json, *fdt, *set.

- The mainly data types used in this project are float number, which is used to record voltage signal, and the timestamp data type, which is used to label the time when start pulsing or recording.

Solution

1. For the first problem, since the whole dataset includes 17 participator, I may just pick one participator's data to put into current project.

2. The dataset of this EEG result includes *.tsv, *json, *fdt, *set. After doing research online, I found this dataset structure is belong to EEGLAB format. On EEGLAB website, they provides EEGLAB software link, which I can download it. However, this EEGLAB package requires MATLAB to run it. Since I have MATLAB software available on my laptop, I can intercept the RAW data in this way.

3. Another solution for reading data or data manipulation is to use MNE package. The MNE package is built based on python environment. And it includes several functions that can importing raw data from different EEG devices.