**Grasp-and-Lift EEG Detection**

<https://www.kaggle.com/c/grasp-and-lift-eeg-detection/data>

1. **The project's domain background** — the field of research where the project is derived

My own personal research interests (machine/deep learning and Virtual Reality) led me to a V.R. programming engineer position at the CINETic Research Center in Bucharest, Romania. I have recently been assigned to our MET Lab which researches EEG (Electroencephalogram) data that represents human brain activity. It has been recorded by measuring the electrical brain activity for epileptic participants which are shown various types of stimuli. We’re trying to implement this paper: <https://www.biorxiv.org/content/10.1101/483487v1> but there’s a pre-processing step that’s being done by one of our collaborators and we’re still waiting for this data to be provided to us.

I digressed a bit from the actual project that I’m proposing for the capstone project because I wanted to give you a bit of context regarding it. It would be a great practice for me to get better familiarised with the type of data and methods that are particular to the field of EEG and that’s why I’m proposing to follow along the lines of the Grasp-and-Lift EEG Detection Kaggle competition (https://www.kaggle.com/c/grasp-and-lift-eeg-detection/data). Some general information about this field can be found here:

<https://en.wikipedia.org/wiki/Electroencephalography>

<https://towardsdatascience.com/using-machine-learning-to-categorise-eeg-signals-from-the-brain-to-words-728aba93b2b3>

<https://www.researchgate.net/publication/262537837_A_Review_on_Machine_Learning_Algorithms_in_Handling_EEG_Artifacts>

<https://www.ncbi.nlm.nih.gov/pubmed/30808014>

1. **A problem statement** — a problem being investigated for which a solution will be defined;

The problem statement is defined in the **Overview** tab of the competition:

This competition challenges you to identify when a hand is **grasping**, **lifting**, and **replacing** an object using EEG data that was taken from healthy subjects as they performed these activities. Better understanding the relationship between **EEG signals** and **hand movements** is critical to developing a **BCI** device that would give patients with neurological disabilities the ability to move through the world with greater autonomy.

1. **The datasets and inputs** — data or inputs being used for the problem;

The dataset description is provided in the **Data** tab of the competition:

There are 12 subjects in total, 10 series of trials for each subject, and approximately 30 trials within each series. The number of trials varies for each series. The training set contains the first 8 series for each subject. The test set contains the 9th and 10th series.

For each **GAL** (grasp-and-lift), you are tasked to detect 6 events:

* HandStart
* FirstDigitTouch
* BothStartLoadPhase
* LiftOff
* Release
* BothReleased

These events always occur in the same order. In the training set, there are two files for each subject + series combination:

* the \*\_data.csv files contain the raw 32 channels EEG data (sampling rate 500Hz)
* the \*\_events.csv files contains the ground truth frame-wise labels for all events

The events files for the test set are not provided and must be predicted. Each timeframe is given a unique id column according to the subject, series, and frame to which it belongs. The six label columns are either zero or one, depending on whether the corresponding event has occurred within ±150ms (±75frames). A perfect submission will predict a probability of one for this entire window.

1. **A solution statement** — the solution proposed for the problem given;

In the preliminary research I did for this project I actually found very different approaches to the given problem. The competition winning solution proposes a complex model ensembling method where a 3 level pipeline feeds data from one type of model to another: <https://www.kaggle.com/c/grasp-and-lift-eeg-detection/discussion/16479>

This is a great starting point but it’s complexity is beyond my knowledge of the domain and of the deep learning field in general so I’ll use it more as a reference guide.

Two approachable solutions with comparable results have been proposed:

a) Hendrik J. Weideman (<https://hjweide.github.io/kaggle-grasp-and-lift>) uses the following network architecture: eight one-dimensional convolutional layers, four max-pooling layers, and three dense layers. The final dense layer has six output neurons, each with a sigmoid activation function that predicts the probability that a given action is active. We use the rectified linear unit (ReLU) as the activation function in all layers except for the output layer. Dropout with p = 0.5 is applied in the first two dense layers.

b) Eszter Schoell at al. (<https://nycdatascience.com/blog/student-works/team-oriented-grasp-and-lift-eeg-kaggle-competition/>) use a method called Common Spatial Pattern (<https://en.wikipedia.org/wiki/Common_spatial_pattern>) to reduce the 32 EEG channels to 4 features which are then fed to a 2 layer neural network.

My intent is to implement one of the two solutions, maybe both and see which one can produce the best results.

1. **A benchmark model** — some simple or historical model or result to compare the defined solution to;

The benchmark model would be in this case the competition winning solution which obtained a 0.98 score

1. **A set of evaluation metrics** — functional representations for how the solution can be measured;

The solution will be measured using the mean of the individual areas under the **ROC** curve for each predicted column (each of the 6 grasp-and-lift tasks).

1. **An outline of the project design** — how the solution will be developed and results obtained.

Like I previously mentioned, I will try to implement at least one of the two solutions proposed at point 4). I will use the standard python libraries that were used throughout this course and the MNE open source library (<https://mne.tools/dev/>) that is specifically designed for EEG data analysis and visualisation.