

Project proposal: Grasp-and-Lift EEG Detection*

Anirudhan Jegannathan Rajagopalan, Michele Cerú†

Abstract. This project aims to classify a human's hand motions from his EEG signal data. This will help in developing Brain-Computer Interface prosthetic devices for restoring a patient's ability to perform basic daily tasks. We are provided with time series EEG recordings of the subjects performing the hand actions which we wish to identify. The current implementation which will be used as baseline uses Convolutional Neural Networks for detecting this with an accuracy of 0.98109 MCAUC. We propose to use Logistic regression with Stochastic Gradient Descent for detecting the hand motions with a performance greater than or equal to the baseline.

1. Introduction. The main goal of this project is to identify hand motions from scalp Electroencephalogram (EEG) recordings, as described in the Kaggle competition[1]. The dataset consists of 3,936 Grasp and Lift (GAL) series, meaning that the analysed subject grasped an object, held it for some seconds and then replaced it (as explained in detail in[2]). Every time, without acknowledging the subject, two main properties of the object were changed: the weight, that could be 165g, 330g or 660g, and the contact surface, that could be sandpaper, suede or silk. In this context there are six events that represents different stages of the hand movements that we aim to predict thorough EEG analysis:

1. *HandStart*: the beginning of the movement.
2. *FirstDigitTouch*: making contact with the object.
3. *BothStartLoadPhase*: starting to load the object.
4. *LiftOff*: holding the object up.
5. *Replace*: replacing the object in its original position.
6. *BothReleased*: releasing the fingers from the object.

The training dataset contains the exact moment when this events occurred during the GAL, that were measured using the 3D position of both hand and object, electromyography signal (EMG) coming from the arm and the hand muscles of the subject, and the force/torque applied to the object. An important restriction to take in account while trying to predict this event, is that for a GAL we can use only data collected in past series and not use the futures one. This restriction is due to the fact that in a real world application there would be no access to future data.

The study aims to find a correlation between the GAL and the EEG that could be applied on developing techniques for the control of prosthetic devices. More in general EEG lay at the base of non invasive brain computer interface BCI[3], that doesn't depends on neuromuscular control and therefore could be used to help patient with heavy neuromuscular disorder to interact with the environment (such as patient who have lost hand function).

2. Performance Criterion. We will use mean column-wise Area Under the Curve (AUC) for evaluating the performance of our output. That is the mean of individual areas under the

*From a Kaggle competition

†New York University (ajr619@nyu.edu; mc3784@nyu.edu). Questions, comments, or corrections to this document may be directed to that email address.

ROC curve for each predicted columns.

3. Problem Formulation.

4. Algorithm. The reference implementation uses

5. Baseline method, algorithm, software. We consider the scripts with most accurate results in the Kaggle competition for baseline[9]. We use the first level implementation by *Cat & Dog*[8] as our baseline. The baseline method uses Logistic regression with LDA for providing an event specific view of the data. There were also two level1 Neural Network approaches that were not event specific.

The model consists of using a covariance matrix for encoding the spatial information between the electrodes. It was found that low frequency data has considerable predictive information. So a “Filter bank” approach was used to give more weightage to low frequency data.

6. short description of the dataset. The datasets consist of data collected with 12 subjects, each of them performing 10 series, each consisting of approximately 30 grasp and Lift performed. The Data is divided in a training set, containing the first 8 series for each subject, and the testing set contains the last two series. In the dataset each of this event correspond to a binary variable (1 if the event is present and 0 otherwise), and we see that the list of event always present in the same other, but the event are not all mutually exclusive meaning that some of them could be 1 at the same time.

REFERENCES

- [1] Kaggle competition: <https://www.kaggle.com/c/grasp-and-lift-eeeg-detection>
- [2] <http://www.nature.com/articles/sdata201447>
- [3] D. J. McFarland and J. R. Wolpaw. Brain-computer interfaces for communication and control. *Commun. ACM*, 54(5):60?66, May 2011.
- [4] <https://github.com/alexandrebarachant/Grasp-and-lift-EEG-challenge>
- [5] <https://hal.inria.fr/hal-00830491/file/journal.pdf>
- [6] <http://www.lacpei.org/LACCEI2010-Peru/published/EInn156.Delgado.pdf>
- [7] Autoregressive Estimation of Short Segment Spectra for Computerized EEG Analysis Jansen, Ben H.Bourne, John R. Ward, James W. Department of Electrical and Biomedical Engineering, School of Engineering, School of Medicine, Vanderbilt University.
- [8] <https://github.com/alexandrebarachant/Grasp-and-lift-EEG-challenge>
- [9] <https://www.kaggle.com/c/grasp-and-lift-eeeg-detection/leaderboard>