SIAM J. APPLIED DYNAMICAL SYSTEMS
Vol. xx, pp. x x–x

Project proposal: Grasp-and-Lift EEG Detection*

Anirudhan Jegannathan Rajagopalan, Michele Cerú†

Abstract. Here abstract to write...

Key words.

AMS subject classifications.

1. Introduction.

The main goal of this project is to identify hand motions from scalp electroencephalogram (EEG) recordings, as described in the kaggle competition that provides the data [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 explain in detail in [2]). Every time, without acknowledging the subject, two main property of the object were changed: the weight, that could be 165, 330 or 660 g, 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 and that we aim to predict though 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 finger 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 the hand and object, electromyography signal (EMG) coming form the arm and the head 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. Meaning that when we analyse a subject performing a GAL we can uses all the data generated in the past ones but not in the future one. This restriction is due to the fact that in real world application there is not 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.

To evaluate the performance on this dataset we will draw the receiver operating characteristic (ROC) curve for each channel of the EEG, and then we will take the mean of the area under the curve (AUC) of all these curves.

*Any comment about who support this work

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

3. problem formulation. (write the mathematical equation)
The first step is to extract features from the data. One possible feature could be the spectral analysis of each channel (as described in in [6]), or the mean Then we will define an M dimensional vector X using all the features as component, to describe each sample of the dataset. Then, given a likelihood function u_λ, that depends on a vector of parameters λ, we can define the score function of a sample as:

$$G_\lambda^X = \nabla_\lambda \log u_\lambda(X) \tag{3.1}$$

This gradient (which dimension depends only on M), represent the direction in which the parameters λ should change in order to fit the model in a better way. Then the similarity between two samples in the dataset X, Y it's given by the Fisher Kernel (as described in [8]):

$$K(X, Y) = G_\lambda^{X'} F_\lambda^{-1} G_\lambda^Y \tag{3.2}$$

where F_λ is the Fisher information matrix:

$$F_\lambda = E_{x\ u_\lambda}[G_\lambda^X G_\lambda^{X'}] \tag{3.3}$$

that could be decomposed as $F_{-\lambda} = L_\lambda^{'} L_\lambda$

4. algorithm. (the one(s) used in the reference papers)

5. baseline method, algorithm, software. The kaggle competition winners [1] provided a solution based on three level. The first level is dedicated to features extraction using mainly two different methods. The first one uses Covariance matrices estimated using a sliding window (of around 500 samples). Considering the six event to predict and the absence of them we can model seven different brain states to predict. They estimated the geometric mean of the covariance matrices, producing in this way a seven dimensional feature vector. The second one is based on the fact that the signal contains predictive information in low frequency, so they used a "Filter Bank" approach to extract this features. Then they combine the results obtained with this two methods to generate the final set of features. The main algorithms used for this level are Logistic Regression, Convolutional Neural network and Recurrent Neural Network.

The second level of the model is trained on the prediction that came from the first level and generate metafeatures that are used in the third level. They applied several different algorithms using bootstrap aggregating to increase the accuracy.

In the third level, predictions from the second level are used to maximise AUC, while avoid overfitting.

6. short description of the real-world datasets. (scales and size of the dataset, missing data, etc.)
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 mutually exclusive meaning that more then one of the can be 1 at the same time.

REFERENCES

[1] Kaggle competition: https://www.kaggle.com/c/grasp-and-lift-eeg-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.laccei.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://hal.inria.fr/hal-00830491/file/journal.pdf