Classfication of EEG signalsl

Carlos Sanchez Ramirez
University of Guelph
50 Stone Road East
Guelph, Ontario, Canada
sanchez@uoguelph.ca

ABSTRACT

Brain Computer Interface (BCI) prosthetic devices have the capability of improving and even enhancing the quality of life of people with limited mobility. Many low-risk, noninvasive BCI devices try to classify EEG signals with the help of different machine learning methods. Discovering which classifiers and features to use is an active area of research as EEG signals are complex signals with spatiotemporal dependencies. In this paper, EEG signals are processed using a common spatial filter and the resulting features are feed to three classifiers, Logistic Regression, Multilayer Perceptron and Random Forest. These will be evaluate based on receiver operating characteristic (ROC) curve and area under the curve (AUC).

1. INTRODUCTION

Brain Computer Interface (BCI) systems provide communication and control capabilities to people with motor disabilities [8]. In cases where people have reduced or limited peripheral muscular activity, BCI systems can be used to restore people's abilities to communicate and perform basic activities of the daily life [6]. These systems generally consist of a signal acquisition hardware, signal processing software and a command devices [8]. When it comes to signal acquisition and processing, many limitations exists in each of these as signals collected are noisy and hard to interpret [6].
Electroencephalogram (EEG) signals are commonly used in BCI systems as it is a non invasive method of measuring brain activity. EEG records electrical of the brain as measured by electrodes placed in the scalp [9]. The brain's electrical activity originates from the activation of individual neurons [9]. Groups a neurons activate in regular patterns that can be analyzed in frequency and domains [3]. Typical EEG recording systems consist of 32 electrodes that capture signals from different areas of the brain [3]. EEG signals are very susceptible to noise that may originate from the electrical activity of the face and eye muscles, and 60 Hz alternating current (AC) line noise [3].
Figure 1 and figure 2 help us better understand EEG signals. Figure 1 shows the electrode placement in a typical EEG recording. Electrodes closer to the face will pick up electrical activity generated by muscles of the face. At any given time, each electrode or channel will record a slightly different signal. Figure 2, shows the shape of the signal over a given period of time. In time domain, the signal appears to be

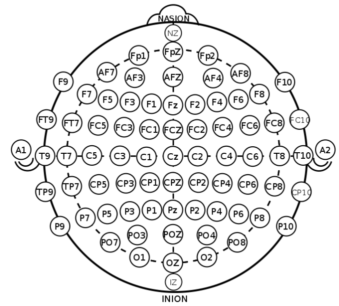

Figure 1: 20-10 Electrode Placement

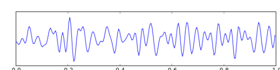

Figure 2: Beta Waves

noisy and hard to interpret which is why frequency domain analysis is commonly performed. EEG signals have distinct frequency bands and each represents different states of brain activity. Beta waves which exist approximately between 12 and 30 Hz are associated with active thinking and are of interest for movement detection applications [9; 3]. When classifying EEG signals, spatiotemporal features should considered.
For this paper, multi-channel EEG recordings of grasp and lift trials will be processed and classified using Logistic Regression, Multilayer Perceptron and Random Forest. The features will be obtained through a common spatial pattern filter. Performance of the three classifiers will be evaluated based on ROC and AUC. The classified signal could then be used to control a computer.

2. METHODS

To classify EEG signals classifications algorithms can be used. In this paper, Logistic Regression, Multilayer Perceptrons and Random Forest will be explored. Before describing the different methods, it is worth describing problems that arise in BCI research. To obtain high classification accuracy the signals need to have low variance; however, EEG signals have high variance across individuals and even within individuals [6]. Another issue that arise is the curse of dimensionality. To accurately describe each of the target classes, the amount of data needed increases exponentially with the dimensionality of the feature vectors [6]. If a large number of features is obtained from a small dataset, the classifier will tend to have low accuracy [6]. In BCI, datasets are typically small and desired feature dimensionality is usually high [6].

2.1 Common Spatial Pattern Filters

Common Spatial Filters(CSP) are commonly used in BCI applications [1]. CSP filters turn multi channel electrode readings and turn it in to a single vector [1].These filters amplify power changes in various frequency bands [1]. This filtering technique is susceptible to overfitting and does not deal well with outliers [1]. In this paper, 4 CSP filters will be used.

2.2 Logistic Regression

Logistic Regression is a linear classifier and involves a probabilistic view of classification [4]. This method is resistant to overfilling, fast to train and makes no assumptions of class distribution but has a linear decision boundary [5]. This classifier has been used by Tomioka et al. in a similar setting as described in their paper [10]. Moreover, this classifier was used as benchmark in a Kaggle competition on which this paper is based on. It was chosen as starting point.

2.3 Multilayer Perceptron

Multilayer Perceptron (MLP) are commonly used in the BCI community as they are very flexible and my adapt to multiple problems [6]. MLP have input and output layers with several hidden layers in between [6]. MLP tend overfit the data which why the architecture has to be carefully selected [6]. MLP are parameterized by the number of hidden neurons, learning rate and activation function.

2.4 Random Forest

Random Forest is ensemble classifier where multiple weak decision trees are trained and together they form a stronger classifier[2]. This method is resilient to noise, performs well and is easy to train as it only has few parameters [2]. It's parameterized by the tree depth and the number of trees used.

3. EXPERIMENT

3.1 Data Description

The data consist of multichannel EEG signals recorded during grasp-and-lift (GAL) trials. The dataset consists of a total of 10 series collected for 12 subjects. For each subject 3,936 GAL trials were collected. 328 trails were collected per series. In each trial, the subjects were asked to reach for the object, grasp it with the thumb and index finger, lift it and hold it for a couple of seconds, put it back on the support surface, release it, and, lastly, to return the hand to a designated rest position [7]. The provided data was labeled as 6 different events. Series 1 through 8 were labeled and 9 and 10 were not. The data was processed to remove electric noise.

3.2 Data Processing

Raw EEG signals were extracted, processed and classified for each subject. Since the signals of interest exist between 7 to 30Hz a bandpass Butterworth filter was applied to the signal. The EEG signal was epoch by using a window of 2 seconds before and after the start of each event. The resulting epoch signal was used to train a common spatial pattern (CSP) filter. A total of 4 CSP filter were trained and regularized using Ledoit-Wolf covariance regularizer. After being trained, the filter was applied to the signal. The signal was then rectified and smoothed by convolving it with a boxcar window. Lastly, a logarithm was applied. The resulting matrix consisted of 4 time series vectors which were used as features to train the models. To cut down on computation cost, the feature vectors were down sampled by a factor of 10. Models were train using series 1-6, validated using series 7 and tested on series 8. Series 9 and 10 were not used as these did not have labels and as result evaluation metrics could have not been calculated.

3.3 Classifier Optimization

After optimizing, the following parameters were chosen. For logistic regression, the inverse of regularization strength (C) was chosen to be 1.0 and it was solved using a liblinear solver. In the case of the MLP, 500 hidden layers with Relu activation function were used with a learning rate of 0.0001 which was solved using stochastic gradient descent. For Random Forest, 100 estimators were trained until every leaf was pure or until the number of minimum split equalled to 2.

4. RESULT

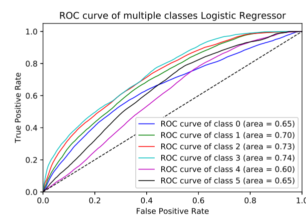

Figure 3: ROC curve obtained using Logistic Regression

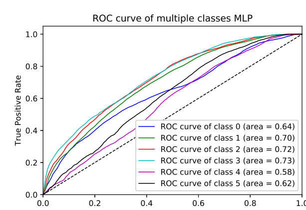

Figure 4: ROC curve obtained using MLP

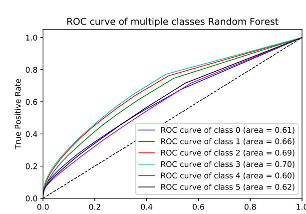

Figure 5: ROC curve obtained using Random Forest

Table 1: Average AUC for the Classifiers

Classifier	Average AUC
Logistic Regression	0.68
MLP	0.66
Random Forest	0.65

4.1 Evaluation

Logistic regression regression produced the highest AUC followed by MLP and Random Forest. It was consistently observed that event 4 had the lowest AUC. This event corresponds to the hand hovering above the object. The signal for this event probably had high variance and the classifiers had trouble dealing with it.

4.2 Limitations

A more systematic approach to tuning the parameters of the algorithms would have improved the AUC but the biggest problem was downsampling the data. Given computational limitations the data had to be down sampled. A total of 36 (3 for each subject) classifiers were implemented, and train trained. This has high computational cost. Cloud computing could have been used but the learning curve associated with it might have taken distracted the objective of the project which was to learn about the data and classifiers.

5. CONCLUSIONS

In this paper, Logistic Regression, MLP and Random Forest were used to classified 6 different events. Data used to train these was multichannel EEG recordings. The signal was filtered using a common spatial filter. Four resulting feature vectors were used to train the classifiers. Logistic Regression obtain the highest AUC. Ultimately, computational cost hindered the ability to finely tune the algorithms.

6. REFERENCES

[1] T. Aksenova, A. Barachant, and S. Bonnet. Spatial filtering optimisation in motor imagery eeg-based bci. *Deuxime confrence franaise de Neurosciences Computationnelles*, Oct 2008.

[2] L. Breiman. Random forest. *Machine Learning*, 45(1):5?32, 2001.

[3] D. R. BRONZINO, JOSEPH D.. PETERSON. *BIOMEDICAL ENGINEERING FUNDAMENTALS*. CRC PRESS, 2017.

[4] H. Dumme. *A Course in Machine Learnin*. Self-Published.

[5] J. Howbert. Machine learning logistic regression.

[6] F. Lotte, M. Congedo, A. Lecuyer, F. Lamarche, and B. Arnaldi. A review of adaptive feature extraction and classification methods for eeg-based brain-computer interfaces. *JOURNAL OF NEURAL ENGINEERING*, 2007.

[7] M. D. Luciw, E. Jarocka, and B. B. Edin. Multi-channel eeg recordings during 3,936 grasp and lift trials with varying weight and friction. *Scientific Data*, 1:140047, 2014.

[8] G. Schalk, D. Mcfarland, T. Hinterberger, N. Birbaumer, and J. Wolpaw. Bci2000: A general-purpose brain-computer interface (bci) system. *IEEE Transactions on Biomedical Engineering*, 51(6):1034?1043, 2004.

[9] M. Teplan. Fundamentals of eeg measurement. *MEASUREMENT SCIENCE REVIEW*, 2(2), 2002.

[10] R. Tomioka, K. Aihara, and K.-R. Muller. Logistic regression for single trial eeg classification.