2017 IEEE 30th International Symposium on Computer-Based Medical Systems

# Detection of Alcoholism based on EEG Signals and Functional Brain Network Features Extraction

Negar Ahmadi, Yulong Pei, Mykola Pechenizkiy
*Department of Mathematics and Computer Science*
*Eindhoven University of Technology (TU/e)*
*Eindhoven, the Netherlands*
*n.ahmadi@tue.nl, y.pei.1@tue.nl, m.pechenizkiy@tue.nl*

***Abstract*—Alcoholism is a common disorder that leads to brain defects and associated cognitive, emotional and behavioural impairments. Finding and extracting discriminative biological markers, which are correlated to healthy brain pattern and alcoholic brain pattern, helps us to utilize automatic methods for detecting and classifying alcoholism. Many brain disorders could be detected by analysing the Electroencephalography (EEG) signals. In this paper, for extracting the required markers we analyse the EEG signals for two groups of alcoholic and control subjects. Then by applying wavelet transform, band-limited EEG signals are decomposed into five frequency sub-bands. Also, the principle component analysis (PCA) is employed to choose the most information carrying channels. By examining various features from different frequency sub-bands, six discriminative features for classification are selected. From functional brain network perspective, the lower synchronization in Beta frequency sub-band and loss of lateralization in Alpha frequency sub-band in alcoholic subjects are observed. Also from signal processing perspective we found that alcoholic subjects have lower values of fractal dimension, energy and entropy compared to control ones. Five different classifiers are used to classify these groups of alcoholic and control subjects that show very high accuracies (more than $90\%$). However, by comparing the performance of different classifiers, SVM, random forest and gradient boosting show the best performances with accuracies near $100\%$. Our study shows that fractal dimension, entropy and energy of channel $C1$ in Alpha frequency sub-band are the more important features for classification.**

***Keywords*-Alcoholism, Brain Network, Classification, EEG, Feature Extraction, Brain Signal Processing.**

## I. INTRODUCTION

Alcoholism, or alcohol use disorder, is a common neurological disorder affecting about $10\%$ of the population [1]. Depending on how much alcohol is taken, drinking alcohol has short-term and long-term effects on the brain. In short-term it may lead to cognitive and mobility impairments such as difficulty in walking, blurred vision, slurred speech, slowed reaction times, impaired memory, upset stomach, blackout and even coma [2]. Binge drinking and continued alcohol use in large amounts are associated with many health problems in long-term such as: alcohol poisoning, high blood pressure, liver disease, nerve damage, gastritis, permanent damage to the brain [3]. Alcoholisms effects may become pronounced by both genetic and environmental factors such as: age, gender, health, and family history [4].

Electroencephalography (EEG) signals are the electrical signals generated in the brain as a result of firing of neurons on the cerebral cortex and scalp. EEG is a very effective tool for studying the complex dynamics of brain activities. Hence, it is used widely to understand the effect of alcoholism on the brain and distinguish alcoholics from normal subjects based on the differences in the EEG signals and brain functions [5].

Frequency-domain and time-domain methods are the most common techniques for assessing alcoholic EEG signals and classifying alcoholic and non-alcoholic subjects [6]. For this purpose, the EEG signal initially undergoes dimensionality reduction if required, followed by feature extraction. The features are then used by the classifier to classify the data as alcoholic and control. There are several studies in the literature working on how EEG differs for alcoholics and non-alcoholics. Power and coherence of various EEG rhythms are an important distinguishing factors for alcoholic and non-alcoholic subjects [5]. It has been shown that alcoholics have increased power of rhythm due to intake of alcohol [7], [8]. Hayden et al. [9] by analysing the frequency power of the EEG signals showed that alcoholism in people may lead to frontal lobe dysfunction and alcoholism subjects exhibit a pattern of frontal asymmetry similar to that found in other psychiatric groups. Winterer et al. [10] reported that in the alcoholic subjects bilateral, intrahemispheric, posterior coherences are significantly increased in the alpha and beta frequency bands both in long-term abstinent and non-abstinent alcohol-dependent subjects. Review of research works on alcoholism diagnosis and prognosis based on the EEG signals are available in refs. [11], [12].

A functional brain network accounts for the neuro-dynamical interactions between neural regions. Functional connectivity defines statistical interdependence between the dynamics of all pairs of the network nodes without taking into account causal effects. The more correlated the activity between two regions, the higher the weight of the functional connections between them [13]. Therefore, the complex network theory is a common statistical physics approach

2372-9198/17 $31.00 © 2017 IEEE
DOI 10.1109/CBMS.2017.46

for investigating the functional brain network. As a result, network features (such as: strength, centrality, and etc.) can be applied in modelling to capture different characteristics of the brain network [14]–[16]. Technically, a feature represents a distinguishing property, a recognizable measurement, and a functional component obtained from a section of a pattern and feature extraction transforms raw signals into more informative signatures or fingerprints of a system that are the most important for classification exercise.

In addition to the network features, the EEG signals also have their own universal features (such as energy, entropy, etc.) that could be suitably used for the brain analysis. Time frequency distributions (TFD), fast Fourier transform (FFT), eigenvector methods (EM), wavelet transform (WT) and auto regressive method (ARM) are the most common techniques to extract the features of EEG signals [17].

In this paper we aim to propose the most discriminative signal/network features at different frequency bands for detecting alcoholism from healthy subjects. By comparing the extracted features, the capability and importance of them for classification purpose are evaluated.

The rest of this paper is structured as follows: The experimental data, decomposition of band-limited EEG to different frequency sub-bands, EEG dimensional reduction, feature extraction and classification are discussed in Methodology section, followed by results and finally conclusions are given in the last section.

## II. Methodology

### A. Wavelet Decomposition

The wavelet is a smooth and quickly vanishing oscillating function with good localization in both frequency and time [18]. Wavelet decomposition has been used to obtain frequency bands of EEG for various problems. The general frequency band of the brain signals is between 0.5-30 Hz. In this study, the wavelet decomposition is used to decompose the EEG signals into five well-known sub-bands, i.e. Delta (0.5-4 Hz), Theta (4-8 Hz), Alpha (8-13 Hz), Beta (13-30 Hz) and finally Gamma ($>$30 Hz)

### B. Dimensional Reduction

The principal component analysis (PCA) is applied to reduce EEG signals information. The PCA is a statistical procedure that uses an orthogonal transformation to convert a set of observations of possibly correlated variables into a set of values of linearly uncorrelated variables called principal components. The number of principal components is less than the number of original variables [20]. The main goal of applying the PCA to an input data set is to have less computational complexity and to improve the accuracy of classification by reducing the feature vectors dimension [21]. Reduction of dimension is necessary for obtaining the most information carrying channel of EEG data set. The PCA can be done by eigenvalue decomposition of a data covariance matrix or singular value decomposition of a data matrix, usually after mean centring and normalizing the data matrix for each attribute. In this study, by using PCA the number of significant EEG channels got reduced from 64 to 12 optimal channels.

### C. EEG Feature Extraction

In this paper, we use different features based on the brain network (i.e. mean synchronization and strength) and EEG signals (fractal dimension, energy and entropy).

*1) Mean synchronization:* For constructing functional brain network, we apply the phase-lag-index (PLI) as a reliable synchronization measure. The alcoholic subjects have impaired synchronization of brain activity. The synchronizations of EEG signals are investigated in all frequency sub-bands. Our analysis shows that the alcoholic brain networks have lower synchronizations in Beta frequency sub-band compared to the control brain networks (see Fig. 1). Therefore, the value of mean synchronization in Beta frequency sub-band is considered as a discriminative feature for classification.

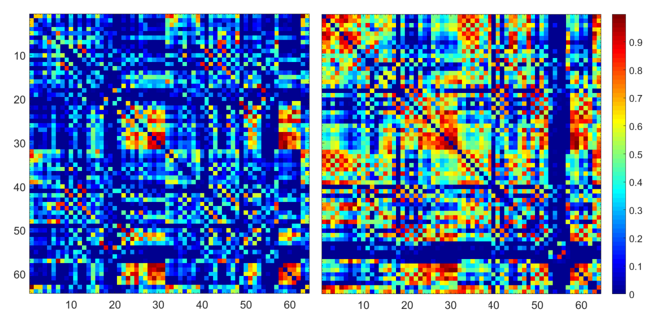



Figure 1: The synchronization matrix of 64 channels in Beta frequency sub-band of an alcoholic (left) and a healthy brain (right).

*2) Strength:* The vertex strength is defined as the sum of weights of links connected to the vertex and is formalized as follows: $S_i = \Sigma w_{ij}$, where $j \in neighbor(i)$ and $w$ represents the weighted adjacency matrix, in which $w_{ij}$ is the weights on the edge between node $i$ and $j$ [23].

Our study shows that the strength of channels associated to left hemisphere of brain in Alpha frequency sub-band is significantly different in alcoholic and control subjects. This result is in line with the lateralization of the brain function that refers to how some neural functions, or cognitive processes tend to be more dominant in one hemisphere than the other. It has been shown that a healthy brain network exhibits strong lateralization compared to alcoholic brain networks as alcoholic networks show loss of lateralization especially in Alpha frequency sub-band [22]. Hence, in order to reduce the complexity of analysis and by considering the difference of strength and lateralization between the left and the right hemispheres of alcoholic and control brains, we choose channels $C1$ and $PO7$ as the most discriminative channels and apply the strengths of them as features for classification.

*3) Fractal dimension:* Fractal dimensions are measures of the self similarity of the signals and their value is usually a non-integer and fractional number, hence this dimension is referred to as fractal [24]. If we consider $X$ as a non-empty compact subset of the real plane, then the capacity dimension is defined as:

$$D_c = \lim_{\epsilon \to 0} \frac{\log N_{min}(\epsilon)}{\log(1/\epsilon)} \tag{1}$$

where $N_{min}(\epsilon)$ is the smallest number of disks with radii $\epsilon$ required to cover $X$. In this work, the fractal dimension is applied to channel $C1$ in Alpha sub-band.

*4) Energy and Entropy:* The energy of signal is defined as [25]

$$E = \Sigma_{i=1}^{n} x_i^2 \tag{2}$$

where, $x_i$ represents the value of signal and $n$ is the total number of samples. Fig. 2 shows band limited EEG signals of a healthy and an alcoholism brain. The calculated value of energy for the healthy and the alcoholism EEG signals are 42380 and 2361 respectively. One can see that alcoholism disorder impairs electrical activity of the brain and leads to less amount of energy compared to a healthy brain. In this work, the energy is calculated in Alpha frequency sub-band for channel $C1$ as a feature for classification detection of alcoholism.

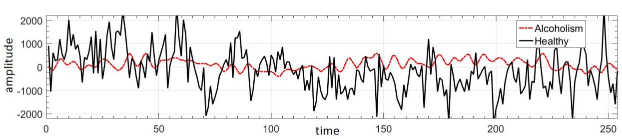



Figure 2: and limited EEG signals for healthy and alcoholism brains.

Entropy measure shows the amount of randomness and uncertainty in the signal, therefore the more fluctuating signal has a higher value of entropy. This measure is defined as:

$$En = \Sigma_{i=1}^{n} x_i^2 \times \log x_i^2. \tag{3}$$

Similar to energy, an alcoholic brain activity gets less complex and shows decreased entropy compared to the healthy brain. Hence, the next marker for finding alcoholism could be the Entropy of EEG signals and we used channel $C1$ entropy in Alpha frequency sub-band as a feature for classification.

### D. Classification

We use 80% instances (80 subjects) as the training data and the rest 20% (20 subjects) as the test data. Since we have limited number of subjects, to avoid over-fitting, we split training and test data randomly and repeat the split process 10 times. Our results are the means and variances of these 10 runs. To explore the importance of features and their combinations in the classification task, we apply 5 widely used classifiers namely: Naive Bayes (NB) [26], support vector machine (SVM) [27], decision tree [28], random forest [29] and gradient boosting [30].

To evaluate the performance of these features in the classification task, two evaluation metrics are applied in the experiments, i.e., *accuracy* and *F1 score*. *Accuracy* counts the number of instances classified correctly and *F1 score* is the harmonic mean of precision and recall. To analyze the classification results in a more comprehensive way, we also apply receiver operating characteristic (ROC) curves [31] in the experiments. ROC curve is a graphical plot that illustrates the performance of a binary classifier system as its discrimination threshold is varied. The curve is created by plotting the true positive ($tp$) rate against the false positive ($fp$) rate at various threshold settings.

## III. Results

The experimental data used in this paper, is a publically available EEG dataset from the University of California, Irvine, the UCI KDD archive [19]. This Multiple electrode time series EEG recordings of control and alcoholic subjects contains measurements from 64 electrodes placed on the scalp with the sampling rate of 256 Hz for one second. The indices of the 64 electrodes are: $Fp1$, $Fp2$, $F7$, $F8$, $AF1$, $AF2$, $Fz$, $F4$, $F3$, $FC6$, $FC5$, $FC2$, $FC1$, $T8$, $T7$, $CZ$, $C3$, $C4$, $CP5$, $CP6$, $CP1$, $CP2$, $P3$, $P4$, $Pz$, $P8$, $P7$, $PO2$, $PO1$, $O2$, $O1$, $X$, $AF7$, $AF8$, $F5$, $F6$, $FT7$, $FT8$, $FPz$, $FC4$, $FC3$, $C6$, $C5$, $F2$, $F1$, $TP8$, $TP7$, $AFz$, $CP3$, $CP4$, $P5$, $P6$, $C1$, $C2$, $PO7$, $PO8$, $FCz$, $POz$, $Oz$, $P2$, $P1$, $CPz$, and, $Y$. Fig. 3 shows the EEG recording positions on brain.

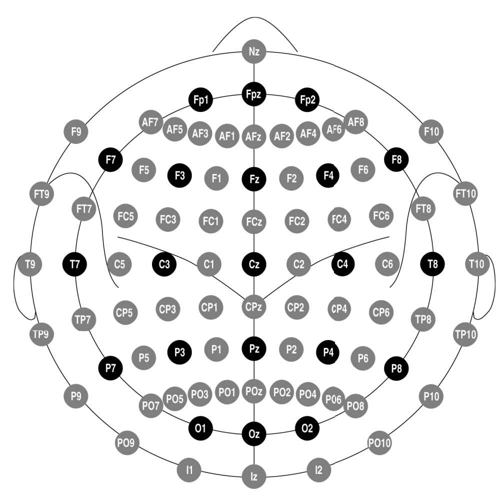



Figure 3: Electrode positions and labels on brain.

Accuracy and F1 score of different feature combinations are shown in Table I and II. Note that "All-X" in these tables denote features without X. BS, ACS, APS, ACEn, ACEt and AFD denote Beta-Sync, Alpha $C1$-Strength, Alpha $PO7$-Strength, Alpha $C1$-Energy, Alpha $C1$-Entropy and Alpha-Fractal Dimension, respectively. From these results, we can draw some conclusions:

- **Overall performance**. All feature combinations perform very well in this classification task since all the accuracy values and F1 scores are larger than 0.9. It indicates the good representative capability of these 6 features. By comparing the performance of different classifiers, SVM, random forest and gradient boosting are the best models and Naive Bayes and decision tree perform worse than them. This observation is consistent with other classification tasks since the first three models are more complicated can better capture the patterns of different classes.
- **Importance of features**. To analyse each feature individually, it can be observed that ACEn, ACEt and AFD are more important than other features because almost all accuracy and F1 score decrease more by removing these three features than removing other features. The rest of these features may be of same importance since by removing either of them the performance changes very little. Furthermore, we use Random Forest and gini importance [32] to analyse feature importances quantitatively. The result is shown in Fig. 6 and it also demonstrates that AFD is the more important feature and followed by ACEt and ACEn.

The ROC curves using different classifiers and all features are shown in Fig. 4. By removing each feature one by one, the ROC curves of different combinations of features are shown in Fig. 5. From these curves, all of the Area under the Curve (AUC) of ROC are larger than 0.9 which indicates good performance of all classifiers in this classification task. By considering the area on each classifier, the conclusion is similar to that in accuracy analysis: SVM, gradient boosting (GB) and random forest (RF) are the best classifiers while Naive Bayes (NB) and decision tree (DT) perform worse than them.

## IV. Conclusion

By analysing EEG signals in both groups of healthy people and subjects with alcohol abuse disorder, the differences between brain electrical activities of these two groups can be detected. Six discriminative features, such as mean synchronization of functional brain network in Beta frequency sub-band, strength of channels C1 and PO7 in left hemisphere in Alpha frequency sub-band, fractal dimension, energy and entropy of channel C1 in Alpha frequency sub-band, were extracted for classification of alcoholic and non-alcoholic subjects. The main effect of alcoholism in functional brain network was abnormality of

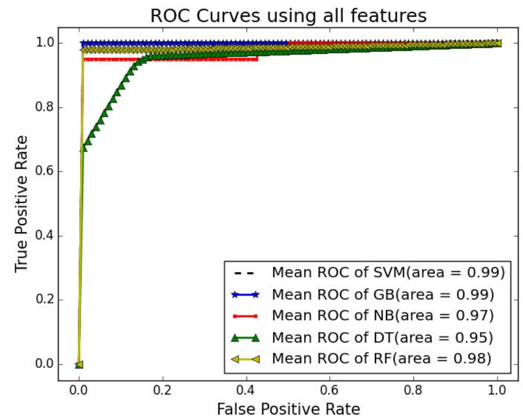



Figure 4: ROC curves using all features.

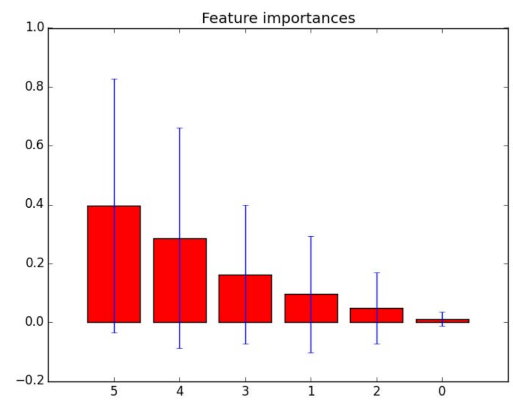



Figure 6: Importances of different features where 0: Beta_Sync, 1: AlphaC1_Strength, 2: AlphaPO7_Strength, 3: AlphaC1_Energy, 4: AlphaC1_Entropy, and 5: Alpha_Fractal D.

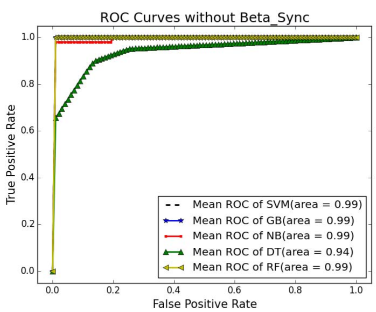
synchronization in left hemisphere of the brain by showing lower Beta frequency sub-band synchronization and loss of lateralization most prominently in Alpha frequency sub-band. From signal processing perspective we found that alcohol abuse can lead to impairment brain electrical activity and causes the lower value of energy, entropy and fractal dimension which are related to the complexity of signals. Naive Bayes (NB), support vector machine (SVM), decision tree, random forest and gradient boosting were used for classification and detection of alcoholism automatically. The performances of all classifiers showed high accuracies over $90\%$. Our results showed that the SVM, random forest and gradient boosting were the best models with accuracy near $100\%$. Also by considering the area under the curves of ROC of each classifiers, the conclusion was similar to that in accuracy analysis. All extracted features performed well in classification tasks and in comparison to each other, the

Table I: Accuracy of different feature combinations in the classification task.

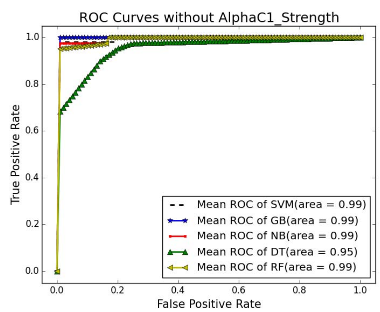
| Feature | NB | SVM | Decision Tree | Random Forest | Gradient Boosting |
|---|---|---|---|---|---|
| All features | 0.96 $\pm$ 0.0044 | 0.98 $\pm$ 0.0016 | 0.94 $\pm$ 0.0044 | 0.99 $\pm$ 0.0009 | 0.98 $\pm$ 0.0016 |
| All - BS | 0.97 $\pm$ 0.0041 | 0.98 $\pm$ 0.0016 | 0.94 $\pm$ 0.0024 | 0.99 $\pm$ 0.0009 | 0.98 $\pm$ 0.0016 |
| All - ACS | 0.96 $\pm$ 0.0044 | 0.98 $\pm$ 0.0016 | 0.94 $\pm$ 0.0064 | 0.98 $\pm$ 0.0016 | 0.98 $\pm$ 0.0016 |
| All - APS | 0.97 $\pm$ 0.0021 | 0.98 $\pm$ 0.0016 | 0.94 $\pm$ 0.0044 | 0.98 $\pm$ 0.0015 | 0.98 $\pm$ 0.0016 |
| All - ACEn | 0.92 $\pm$ 0.0056 | 0.96 $\pm$ 0.0024 | 0.93 $\pm$ 0.0061 | 0.97 $\pm$ 0.0041 | 0.97 $\pm$ 0.0021 |
| All - ACEt | 0.93 $\pm$ 0.0041 | 0.96 $\pm$ 0.0024 | 0.93 $\pm$ 0.0101 | 0.97 $\pm$ 0.0032 | 0.96 $\pm$ 0.0044 |
| All - AFD | 0.95 $\pm$ 0.0044 | 0.98 $\pm$ 0.0016 | 0.92 $\pm$ 0.0095 | 0.95 $\pm$ 0.0026 | 0.97 $\pm$ 0.0041 |

Table II: F1 score of different feature combinations in the classification task.

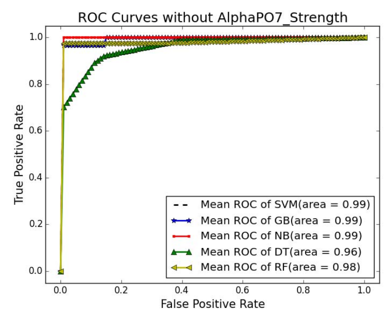
| Feature | NB | SVM | Decision Tree | Random Forest | Gradient Boosting |
|---|---|---|---|---|---|
| All features | 0.9609 $\pm$ 0.0012 | 0.9872 $\pm$ 0.0007 | 0.9389 $\pm$ 0.0131 | 0.9791 $\pm$ 0.0012 | 0.9809 $\pm$ 0.0015 |
| All - BS | 0.9548 $\pm$ 0.0058 | 0.9847 $\pm$ 0.0010 | 0.9505 $\pm$ 0.0045 | 0.9870 $\pm$ 0.0008 | 0.9832 $\pm$ 0.0016 |
| All - ACS | 0.9653 $\pm$ 0.0023 | 0.9842 $\pm$ 0.0011 | 0.9449 $\pm$ 0.0063 | 0.9646 $\pm$ 0.0054 | 0.9823 $\pm$ 0.0013 |
| All - APS | 0.9665 $\pm$ 0.0032 | 0.9831 $\pm$ 0.0016 | 0.9598 $\pm$ 0.0016 | 0.9825 $\pm$ 0.0016 | 0.9756 $\pm$ 0.0029 |
| All - ACEn | 0.9466 $\pm$ 0.0009 | 0.9524 $\pm$ 0.0056 | 0.9480 $\pm$ 0.0026 | 0.9676 $\pm$ 0.0031 | 0.9776 $\pm$ 0.0020 |
| All - ACEt | 0.9405 $\pm$ 0.0057 | 0.9575 $\pm$ 0.0031 | 0.9426 $\pm$ 0.0075 | 0.9591 $\pm$ 0.0060 | 0.9609 $\pm$ 0.0043 |
| All - AFD | 0.9521 $\pm$ 0.0060 | 0.9762 $\pm$ 0.0028 | 0.9452 $\pm$ 0.0040 | 0.9555 $\pm$ 0.0048 | 0.9746 $\pm$ 0.0017 |



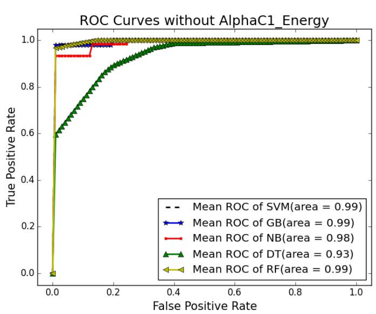

(a) ROC without Beta_Sync



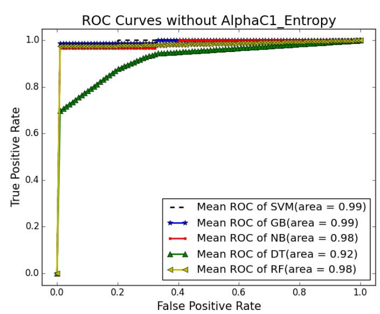

(b) ROC without AlphaC1_Strength



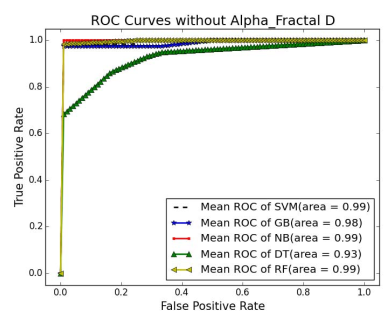

(c) ROC without AlphaPO7_Strength




(d) ROC without AlphaC1_Energy




(e) ROC without AlphaC1_Entropy




(f) ROC without AlphaC1_Fractal D

Figure 5: ROC Curves using different combinations of features.

fractal dimension, entropy and energy of the signals were the most discriminative features because all accuracy values and F1 scores decrease more by removing these three features compared to other features removing and again the results of ROC analysis were similar and showed the importance of those three features for classification.

## References


[1] World Health Organization, *Global status report on alcohol and health*, World Health Organization, 2014.

[2] K.M. Jennison, The shortterm effects and unintended longterm consequences of binge drinking in college: a 10year followup study, *Am. J. Drug. Acohol. Abuse.*, 30(3), 659-684, 2004.

[3] K.J. Sher, B.D. Bartholow, S. Nanda, S., Short-and long-term effects of fraternity and sorority membership on heavy drinking: a social norms perspective, *Psychol. Addict. Behav.*, 15(1), 42, 2001.

[4] M. Wankerl, et al. Effects of genetic and early environmental risk factors for depression on serotonin transporter expression and methylation profiles, *Transl. Psychiatry*, 4(6), e402, 2014.

[5] M. Rangaswamy, et al. Beta power in the EEG of alcoholics, *Biol. Psychiatry.*, 52(8), 831-842, 2002.

[6] D.P. Subha, et al., EEG signal analysis: a survey. *J. med. syst.*, 34(2), 195-212, 2010..

[7] K.E. Thornton,D.P. Carmody, Efficacy of traumatic brain injury rehabilitation: Interventions of QEEG-guided biofeedback, computers, strategies, and medications, *Appl. Psychophysiol. Biofeedback*, 33(2), 101, 2008.

[8] C.L. Ehlers, et al., EEG alpha and level of response to alcohol in Hispanic-and non-Hispanic-American young adults with a family history of alcoholism, *J. stud. alcohol*, 65(3), 301-308, 2004.

[9] E.P. Hayden, Elizabeth, et al., Patterns of Regional Brain Activity in AlcoholDependent Subjects, *Alcohol. Clin. Exp. Res.*, 30(12), 1986-1991, 2006.

[10] G. Winterer, et al., EEG phenotype in alcoholism: increased coherence in the depressive subtype, *Acta Psychiat. Scand.*, 108(1), 51-60, 2003.

[11] U.R. Acharya, et al., Computer-aided diagnosis of alcoholism-related EEG signals, *Epilepsy Behav.*, 41, 257-263, 2014.

[12] B. Tabakoff, P.L. Hoffman, The neurobiology of alcohol consumption and alcoholism: an integrative history, *Pharmacol. Biochem. Behav.*, 113, 20-37, 2013.

[13] D. Papo, et al., *Complex network theory and the brain*, The Royal Society, (2014).

[14] R. Mikail, O. Sporns, Complex network measures of brain connectivity: uses and interpretations, *Neuroimage*, 52(3), 1059-1069, 2010.

[15] G. Prasad, et al., Brain connectivity and novel network measures for Alzheimer's disease classification, *Neurobiol. Aging*, 36, S121-S131, 2015.

[16] N. Ahmadi, M. Pechenizkiy, Application of Horizontal Visibility Graph as a Robust Measure of Neurophysiological Signals Synchrony, *Computer-Based Medical Systems (CBMS), 2016 IEEE 29th International Symposium on*, IEEE, 273-278, 2016.

[17] B. Hosseinifard, et al., Classifying depression patients and normal subjects using machine learning techniques and nonlinear features from EEG signal, *Comput. Methods Prog. Biomed.*, 109(3), 339-345, 2013.

[18] H. Adeli, Z. Zhou, N. Dadmehr, Analysis of EEG records in an epileptic patient using wavelet transform, emphJ. Neurosci. Methods, 123(1), 69-87, 2003.

[19] K. Bache, M. Lichman, *UCI machine learning repository*, 2013.

[20] S. Lipovetsky, PCA and SVD with nonnegative loadings, *Pattern Recognition*, 42(1), 68-76, 2009.

[21] H. Zou, T. Hastie, R. Tibshirani, Sparse principal component analysis, *J. Comp. Graph. Stat.*, 15(2), 265-286, 2006.

[22] C.A. Hanlon, et al., Loss of laterality in chronic cocaine users: an fMRI investigation of sensorimotor control, *Psychiatry Res. Neuroimaging*, 181(1), 15-23, 2010.

[23] P. Van Mieghem, *Graph spectra for complex networks*, Cambridge University Press, 2010.

[24] R. Uthayakumar, Fractal dimension in Epileptic EEG signal analysis, *Applications of Chaos and Nonlinear Dynamics in Science and Engineering*, Vol. 3, Springer Berlin Heidelberg, 103-157, 2013.

[25] D. Gajic, et al., Classification of EEG signals for detection of epileptic seizures based on wavelets and statistical pattern recognition, Biomed Eng-App. Bas. C., 26, 1450021, 2014.

[26] K.P. Murphy, *Naive bayes classifiers*, University of British Columbia, 2006.

[27] C. Cortes, V. Vapnik, Support-vector networks, *Machine learning*, 20(3), 273-297, 1995.

[28] S.R. Safavian, DA survey of decision tree classifier methodology, *IEEE Trans. Syst. Man Cybern. Syst.*, 21(3), 660-674, 1991.

[29] X.W. Chen, L. Mei, Prediction of proteinprotein interactions using random decision forest framework, *Bioinformatics* 21, 4394-4400, 2005.

[30] L. Mason, et al., Boosting Algorithms as Gradient Descent, *NIPS*, 512-518, 1999.

[31] T. Fawcett, An introduction to ROC analysis, *Pattern recognition letters*, 27(8), 861–874, 2006.

[32] L. Breiman et al., *Classification and regression trees*, CRC press, 1984.