



Comparison between Alcoholic and Control Subjects in EEG signals Using Classification Methods

Shaymaa Adnan Abdulrahman

PHD, candidate in Ain Shams University, Egypt

Lecturer at faculty of information technology, imam ja'afar al-sadiq university
Iraq, Bagdad

Corresponding Author Email: shaymaaa416@gmail.com

<https://doi.org/10.26782/jmcms.2019.08.00008>

Abstract

Alcoholism could be identified through analyzing electroencephalogram (EEG) signals. Yet, it is difficult to analyze with multi-channel EEG signal since it is frequently needing long time for execution and complex calculations. The presented paper proposed 13 optimal channel to feature extraction. Firstly, 1200 recordings of biomedical signals will be presented for extracting the sample entropy. Statistical analysis approach will be utilized for the purpose of choosing the best channels for identifying abnormalities in alcoholics. Secondly four classifiers are applied at the decision level, Naïve Bayes, SVM, Logistic Regression, KNN, the accuracy was 80.1%, 92.5%, 73.7% and 90.3% respectively, in this study the SVM classifier is more accuracy.

Keyword : EEG signal, optimal channel, abnormalities in alcoholics, SVM classifier

I. Introduction

Biometrics is the operation of uniquely determined individuals depend on one or more physical, behavioral or cognitive characteristics[XVII].

Physiological biometrics is refer to the physical characteristics of one or more body sections like (iris , face, or hand geometry but behavioral biometrics using information extract from person's behavior, such as typing rhythm, gait and signature. Cognitive biometrics depend on the outputs of the central and peripheral nervous systems, such as the electroencephalogram (EEG), electrocardiogram (ECG), and sudation patterns [V].

Various action types could impact an electrical action in the brain, such as fantasy, moving leg or arm, visualization, or just solving a problem. Electrodes placed over the scalp are used for the purpose of assessing the brain's electrical activity.

The signals of the brain could be acquired through the use of different methods. It could be classified as non-invasive and invasive method. A surgical intervention is

*Copyright reserved © J. Mech. Cont.& Math. Sci.
Shaymaa Adnan Abdulrahman*

needed in the invasive approach to put the electrodes under the scalp. Because of medical hazards and researchers have a tendency to avoid invasive method. While when using noninvasive method, the electrodes are placed on the scalp of human [II].

alcoholism the important of BCI application for medical lies detection/ prevention in the probable defeat of job and the reduction of attentiveness level caused by drinking alcohol and smoking, researcher try to exposed the upper reacting brain part to alcoholism [XVIII]. Alcoholism can be defined as a general neurological disease resulting from mutual impact of environment and genetic aspects. It not just loses the brain system but in addition, it might cause mobility and cognitive weakness [XIII]

II. Related work

The authors of [XIX] proposed when subjected to different action conditions, the controls and subjects alcoholics based on cognitive demand, they encounter non-matching delay task condition. The cognitive demand is reflected on the features of EEG. Signal source separate over a dense electrode system is implemented via the use of Independent Component Analysis (ICA) algorithm. [IX] suggested 3 test situations applied with distinct classifiers. The first situation includes using raw data to be the feature set, the second situation, obtained the features following wavelet decomposition such as entropy, energy, inter-quartile range and median absolute deviation regarding all used sub bands. The last situation includes using the raw data and derived characteristics referred to as hybrid feature set for classification. [VIII] The suggested approach to classify non-alcoholic and alcoholic individuals through the use of EEG signal include 3 main stages: filtering, high pass IIR filter with zero phase distortion is utilized, feature extraction, reflection coefficients regarding filtered EEG signal will be extracted through the use of autocorrelation values in recursive fashion, are suggested as feature. Classification, K nearest neighbor (KNN) classifier is utilized in leave one out cross validation approach. [III] the principle component analysis (PCA) is utilized for choosing the major information that carry the channels. Through analyzing various features from distinct frequency sub-bands, 6 discriminative features for classification have been selected. [XX] for the purpose of detecting brain abnormalities in automatic way. used Wavelet Packet Decomposition (WPD) approach for extracting features, PCA for reducing dimensions, and Back Propagation NN enhanced with GA for classifying alcohol addiction. [XV] the records of EEG are usually disturbed through noise like heartbeat, muscle movement and eye blinking. suggested, Independent Component Analysis (ICA), as noise removal, Stationary Wavelet Transform (SWT) as a feature extraction procedure and are categorized to 2 categories, which are, alcoholism and normal through the use of Probabilistic Neural Network (PNN). [XIV] suggested a new relative wavelet bi-spectrum (RWB) method for EEG signal feature extraction approach in order to distinguish the signal between non-alcoholic and alcohol individuals. discrete wavelet transformation (DWT) is execution substituting the FFT that commonly is utilized in calculating bi-spectrum. Utilizing cross validation, the highest results from RWB feature extraction approach with NN classifier reached approximately 90% rate of recognition.

III. Proposed approach

Step 1: Experimental data that is utilized in the presented study have been acquired from University of California, Irvine Knowledge Discovery in Data-bases Archive UCI KDD [IV]. The dataset have been collected from 122 individuals. All individuals completed 120 trials with 3 types of stimuli [XXIII]. The recordings taken from each individual consist of 61 channel EEG signals, 2 EOG channels and single reference electrode. Sampling rate regarding all channel data is 256 Hz, and duration regarding every trial is single second. There are 3 datasets that are respectively, SMNI_CMI_TRAIN, SMNI_CMI_TEST and FULL. In the presented research, just the first 2 databases are utilized since FULL datasets consist of a small number of all-zero recordings [XXIV]

SMNI_CMI_TRAIN include six hundred recorded files, with each recording consist of signals from 64 electrodes caps. The indices of 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”, “nd” and “Y”. Electrodes X and Y are considered to be EOG signals, while nd are reference electrodes. The nd will be eliminated in this analysis. Therefore, features will be obtained from 63 channels. As indicated in figure 1

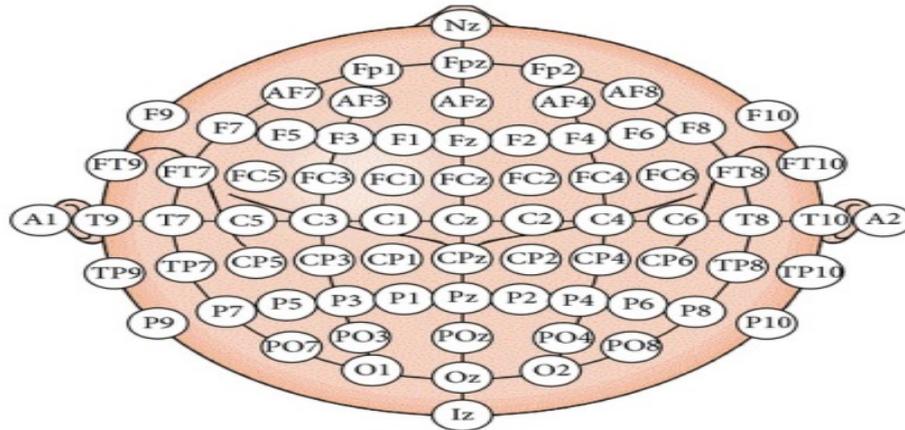


Figure 1 system of electrode placement

In this paper using 13 channel like(FPZ,APZ,FZ,FCz,Cz, C1,C2,C3, C4,FC1,FC2,FC3,FC4) this channel refer to optimal channel. The block diagram representation is shown in figure2

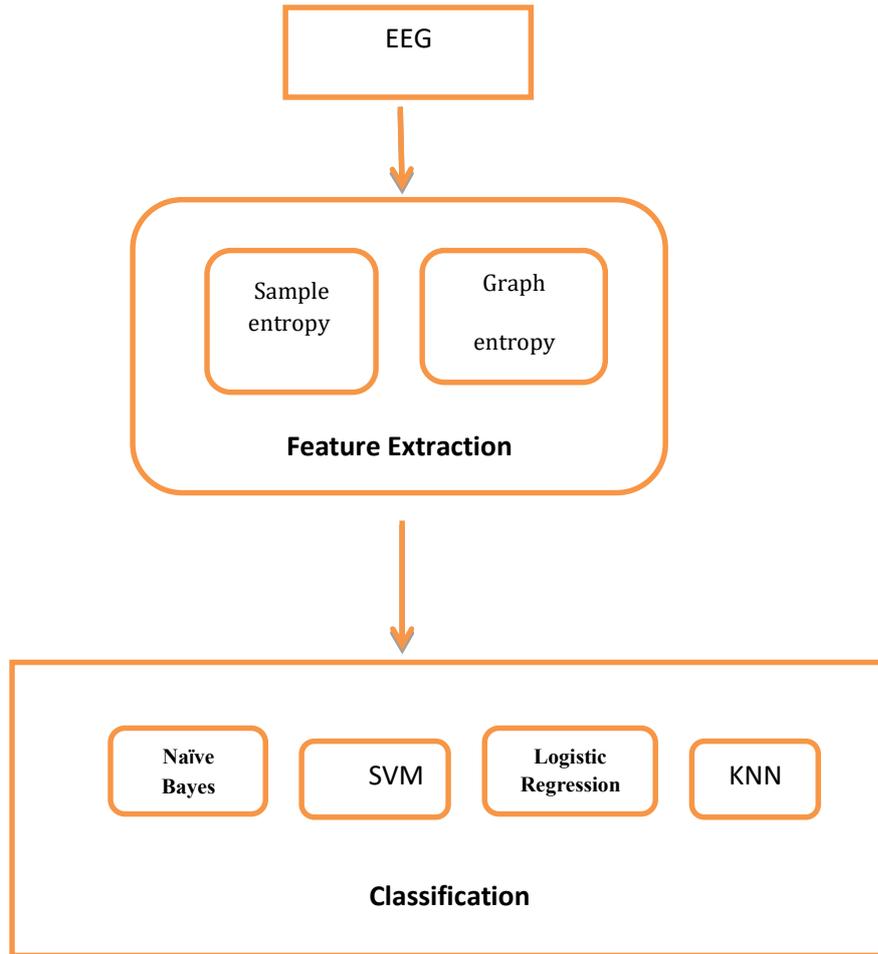


Figure 2 diagram of proposed search

Feature Extraction

Sample entropy

There are many types of graph entropy calculation approaches according to edges or vertex [17]. The presented paper knows graph entropy (GE) with the formula of Shannon's entropy [XXI]. show figure 2

$$h = - \sum_{k=1}^n p(k) \log(p(k)) \tag{1}$$

Sample entropy has been suggested via Moorman and Richman [XVI]. It was utilized for the purpose of measuring complexity of alcoholic EEG [XXII]. The algorithm of sample entropy was applied in the presented paper for estimating the Sample entropy that is available in Physione web-site (<http://www.physionet.org/physiotools/sampen/>). There are 3 input parameters in sample entropy algorithm (1) m: embedded dimension, (2) r: similarity criterion, (3) n: time series length. As displayed in figure 3

Sample entropy could be specified in the equation 2

$$\text{Sample Entropy (m,r)} = \ln \left[\frac{B^m(r)}{A^m(r)} \right] \tag{2}$$

Where

$$A^m(r) = \frac{1}{N-m} \sum_{k=1}^{N-m} A_k^m(r)$$

$$B^m(r) = \frac{1}{N-m} \sum_{k=1}^{N-m} B_k^m(r) \tag{3}$$

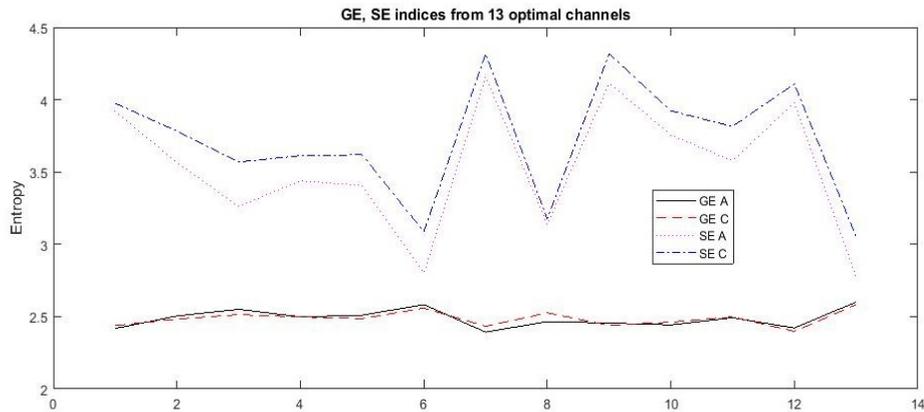


Figure 3 sample entropy and graph entropy with 13 channel of alcoholic and control subject

Classifiers

After that, the computed features will be fed to the classifier for the purpose of classification between different states of the brain of human. We utilized these classifiers in our study.

Naïve Bayes: It depends on the Bayes theory, also it does make a statement that all the attributes of certain class is independent of values of other attributes. Class conditional independence depends on this statement.

SVM classifier: SVMs are the optimum state of the art classifier which have low in comparison to other classifiers such as fuzzy classifiers and neural networks. SVMs depends on finding hyper plane that have the ability of classifying data to distinct classes with a potential for maximum margin. Linear SVMs could classify the non-linear classification task through the use of kernel trick with limited increase in the complexity. There are various benefits in the classifiers of SVM, such as., less parameters tuned in manual way, insensitive to overtraining, margin maximization, and so on.

Logistic Regression: That depends on statistical modeling approach. For example, probability P1 of any dichotomous outcome event could be associated to self-explanatory variables of particular form in equation

$$\text{Logit}(p) = \beta_0 + \sum_{i=1}^n \beta_i x_i$$

Where β_0 is called intercept and $\beta_1, \beta_2, \beta_3 \dots \dots \dots \beta_n$ are the coefficients that are related with certain explanatory variables $x_1, x_2 \dots \dots \dots x_n$

K-Nearest Neighbour (K-NN): Can be considered as an algorithm that is chosen for conducting binary classification, the algorithm is conventional pattern recognition approach, that is statistical supervised classification. The main thought is that certain new test data t, algorithm acquires the K-NN from training set depending on distance between training set and t. The most dominated class among the K neighbours is allocated as class of t as displayed in figure 4 refers to c KNN and SVM as classifier. The classification accuracy acquired with a lot of processes are showed in Table I. From the presented table, SVMs offer optimum results in comparison to KNN, Logistic Regression, and Naive Bayes.

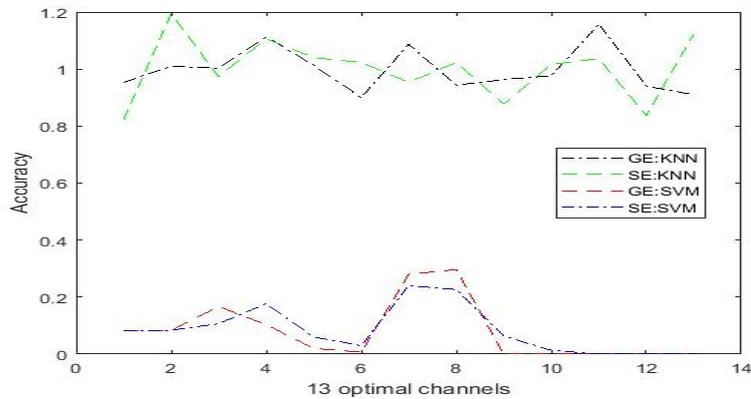


Figure 4 classification with SVM and KNN

TABLE 1 Classification Results for alcoholic EEG signals

Classifier type	# of channels	Accuracy
Naïve Bayes	13	80.1%
SVM	13	92.5%
Logistic Regression	13	73.7%
KNN	13	90.3%

IV. Conclusions

The signals of the brain imitate the control behavior and handled actions of the brain or the impact of collected information from other parts of the body either sensing or internal organs. In this paper, we used UCI KDD as data set . we applied nonlinear features e.g., Sample entropy and , the graph entropy (GE) with Shannon’s entropy are computed features have the ability of classifying between control and alcoholic subject of human brain. the Naiva Bayes , SVM, Logistic Regression and KNN as classifier , SVM classifier produced much better results in comparison to other classifiers.

V. Acknowledgments

The authors show their huge gratitude to Prof. Henri Begleiter at the Neurodynamics Laboratory at the State University of New York Health Centre at Brooklyn, USA whose group documented the raw VEP data and Mr. Paul Conlon, of Sasco Hill Research, USA for EEG data sharing.

*Copyright reserved © J. Mech. Cont.& Math. Sci.
Shaymaa Adnan Abdulrahman*

References

- I. Andreas H, Bernhard O, Christof S, Valdez AC, Schaar AK, Ziefle M, Dehmer M (2013) On graph entropy measures for knowledge discovery from publication network data. In: Cuzzocrea A, Kittl C, Simos DE, Weippl E, Xu L (eds) Availability, reliability, and security in information systems and HCI, pp 354–362. Springer, Berlin.
- II. Abdulkader, S.N., Atia, A., Mostafa, M.S.M.: Brain computer interfacing: applications and challenges. *Egypt. Inf. J.* 16, 213–230 (2015).
- III. Ahmadi N., Pei Y., Pechenizkiy M., Detection of Alcoholism based on EEG Signals and Functional Brain Network Features Extraction , 2017 IEEE 30th International Symposium on Computer-Based Medical Systems.
- IV. Bache K, Lichman M (2013) UCI machine learning repository. University of California, Irvine, School of Information and Computer. <http://archive.ics.uci.edu/ml>.
- V. Cempírek, M. - Šťastný, J. The optimization of the EEG-based biometric classification. *Applied Electronics*. 2007, pp. 25-28.
- VI. Dehmer M, Mowshowitz A (2011) A history of graph entropy measures. *Inf Sci* 181(1):57–78.
- VII. Dehmer M, Mowshowitz A (2011) A history of graph entropy measures. *Inf Sci* 181(1):57–78.
- VIII. Fattah S. A., Fatima K., and Shahnaz C., An Approach for Classifying Alcoholic and Non- Alcoholic Persons Based on Time Domain Features Extracted From EEG Signals, 2015 IEEE International WIE Conference on Electrical and Computer Engineering.
- IX. Gopika Gopan K, Neelam Sinhay and Dinesh Babu, Hybrid Features based Classification of Alcoholic and Non-alcoholic EEG, International Institute of Information Technology, 978-1-4799-9985-9/15/\$31.00 ©2015 IEEE.
- X. Hofeld T, Burger V, Hinrichsen H, Hirth M, Tran-Gia P (2014) On the computation of entropy production in stationary social networks. *Soc Netw Anal Min* 4(1):1–19.
- XI. Hyperparameter. <https://en.wikipedia.org/wiki/Hyperparameter>
- XII. Luque B, Lacasa L, Ballesteros F, Luque J (2009) Horizontal visibility graphs: exact results for random time series. *Phys Rev E* 80(4):046103.
- XIII. Oscar-Berman M, Marinkovi K (2007) Alcohol: effects on neurobehavioral functions and the brain. *Neuropsychol Rev* 17(3):239–257.

- XIV. Purnamasari P. D., Ratna A. A. P., Kusumoputro B., Relative Wavelet Bispectrum Feature for Alcoholic EEG Signal Classification Using Artificial Neural Network, 978-602-50431-1-6/17/\$31.00 ©2017 IEEE.
- XV. Rachman N.T., Tjandrasa H., Faticah C., Alcoholism Classification based on EEG Data using Independent Component Analysis (ICA), Wavelet De-noising and Probabilistic Neural Network (PNN), 2016 International Seminar on Intelligent Technology and Its Application, 2016 IEEE.
- XVI. Richman JS, Moorman JR (2000) Physiological time-series analysis using approximate entropy and sample entropy. *Am J Physiol-Heart Circulat Physiol* 278(6):H2039–H2049.
- XVII. Smit D. J. A. and Posthuma D., Boomsma D. I. and De Geus E. J. C., Heritability of background EEG across the power spectrum *Psychophysiology Journal*, vol. 42, 2005, pp.691-697.
- XVIII. Shooshtari MA, Setarehdan SK Selection of optimal eeg channels for classification of signals correlated with alcoholabusers In: Signal Processing [ICSP].2010 IEEE 10th international Conference on .IEEE: 2010. P.1-4
- XIX. Singh S. N., Godiyal A. K., Panigrahi B.K., Anand S., Source Localization in Alcoholic and Control Subjects to Estimate Cognitive load using EEG Signal. IEEE International Conference on Computer, Communication and Control (IC4-2015).
- XX. Saddam M., Tjandrasa H., Navastara D. A., Classification of Alcoholic EEG Using Wavelet Packet Decomposition, Principal Component Analysis, and Combination of Genetic Algorithm and Neural Network, 2017 IEEE International Conference on Information & Communication Technology and System.
- XXI. Shannon CE (2001) A mathematical theory of communication. *ACM SIGMOBILE Mob Comput Commun Rev* 5(1):3–55.
- XXII. Zou Y, Miao D, Wang D (2010) Research on sample entropy of alcoholic and normal people. *Chin J Biomed Eng* 29:939–942.
- XXIII. Zhang XL, Begleiter H, Porjesz B, Litke A (1997) Electrophysiological evidence of memory impairment in alcoholic patients. *Biol Psychiatr* 42(12):1157–1171
- XXIV. Zhu G, Li Y, Wen P (2011) Evaluating functional connectivity in alcoholics based on maximal weight matching. *J Adv Computat Intell Intell Inform* 15(9):1221–1227