



Improving Alcoholism Diagnosis: Comparing Instance-Based Classifiers Against Neural Networks for Classifying EEG Signal

Shelia Rahman¹ , Tanusree Sharma² , and Mufti Mahmud³ 

¹ Institute of Information Technology, Jahangirnagar University, Dhaka, Bangladesh
sheliatulyar@gmail.com

² University of Illinois at Urbana-Champaign, Champaign, IL, USA
tsharma6@illinois.edu

³ Department of Computing and Technology, Nottingham Trent University,
Nottingham NG11 8NS, UK
mufti.mahmud@ntu.ac.uk, muftimahmud@gmail.com

Abstract. Alcoholism involves psychological and biological components where multiple risk factors come into play. Assessment of the psychiatric emergency is a challenging issue for clinicians working with alcohol-dependent patients. Identifying alcoholics from healthy controls from their EEG signals can be effective in this scenario. In this research, we have applied two instance-based classifiers and three neural network classifier to classify Electroencephalogram data of alcoholics and normal person. For data preprocessing, we have applied discrete wavelet transform, Principal component analysis and Independent component analysis. After successful implementation of the classifiers, an accuracy of 95% is received with Bidirectional Long Short-Term Memory. Finally, comparing the performance of the two categories of algorithms, we have found that neural networks have higher potentiality against instance-based classifiers in the classification of EEG signals of alcoholics.

Keywords: Alcoholism · EEG signal · Machine learning · Instance based classifier · Neural networks

1 Introduction

Alcoholism can be the cause of depression, anxiety, domestic violence, psychosis, and antisocial behavior and in extreme cases, psychiatric disorders. Brains are one of the commonly affected organs in alcoholism causing cognitive, emotional and behavioral disorders. Our brain is a complex system comprised of millions of interconnected neurons that controls our functional and cognitive activity by passing electrical signals among themselves. Alcoholism can damage brain cells consequently changing the electrical activity responsible for brain function. Electroencephalogram (EEG) is a popular medical test and can be used

to detect the abnormalities caused due to alcoholism by analyzing the electric signals recorded. EEG signals are very small non-stationary, nonlinear electrical signals measured only in micro-volts (μV) also tends to change from subject to subject. For a naked eye observation, EEG signals from both alcoholic and healthy control may look nothing different. The small variations are hard to detect for a physician just by looking at it. When a stimulus is presented to test subject, his/her brain might produce some neural responses against the stimulus. In EEG this neural response from a number of similar neurons are recorded in μV . These responses can differ from brain regions to regions. Machine learning (ML) algorithms can be useful for EEG processing to represent underlying frequency structure with its' mathematical models and classification models of neural response values. These models use neural response recordings as attributes to predict if a subject is different from another one. ML models uses mathematical and statistical approaches which enables classification procedure to be faster and more accurate to predict if a subject is alcoholic or not. This analysis can be helpful to build brain computer interface (BCI) for physicians for faster and accurate identification of alcoholic patients.

The aim of our research is two folds. In this research we will observe how well the ML algorithms can predict alcoholism. In addition, we compare the performance two different types of algorithms: Instance based learning and Neural Network. The reason for categorizing the algorithms is EEG signals are mainly time series data and we aim to observe if any particular category of ML has better performance than other one in analysing them. In most researches, new classification models have been proposed or used to predict alcoholism. However, in this study we observe the classification credibility of two different groups of algorithms as well as compare them in regards of time series EEG data. BCI is a great medium for building interpretive models to study behavior of our brains. Determining the best ML for designing a BCI is challenging. With the increase in alcohol consumption rate alcoholism is becoming a serious health and social issue. However, alcoholism is different from a typical drinking habit. Therefore, an early and accurate prediction is highly anticipated. Analysing EEG signals by ML algorithms, this work puts forward such prediction to facilitate the detection of minor abnormality in the brain signals which is nearly impossible for a physician just with naked eye observation.

The rest of the paper is organized as follows– Section 2 talks about the background of the work, and Sect. 3 discusses materials and methods. Section 4 provides experiment results including comparison among results of ML algorithms. Finally, concluding remarks and possible future works are depicted in Sect. 5.

2 Background

Alcoholism, also known as alcohol consumption disorder (AUD) refers to a condition of alcohol abuse where brain is one of the commonly affected organs. Alcoholics are reported to have less cortical grey, white matter volumes as well as reduced volumes of sulcal and ventricular CSF (cerebrospinal fluid) when compared to non-alcoholics [6]. The disturbance in the functional connectivity can be

detected by recording the electrical signals in the brain. EEG records electrical activities in brain using electrodes. Contrast to the commonly questionnaire-based alcoholism identification methods, EEG represents changes of biophysical response in the cerebral cortex offering more accurate diagnosis of alcoholism. However, EEG signal itself is very random in nature and computation complex. ML algorithms provides mathematical models that automatically identifies the underlying structure in the EEG signal to identify the distinct frequency level responsible for different brain activity thus differentiating abnormal brain condition from healthy controls.

Numerous studies have been conducted regarding classification of abnormal subjects from healthy ones using ML on EEG data. ML provides an automated method with adaptability and generalization capability [11, 12] enabling analysis of complex EEG signals offering less human intervention. Some promising examples are CNN (convolutional neural network) for detection of Parkinson's disease from EEG data [14, 16]; experiment on emotion recognition [10] using K-nearest neighbor (KNN) classifier from EEG data. Automatic seizure detection from EEG data using ML has proven to be very successful in previous study [17]. Similar to these abnormal brain conditions, ML is tend to be used on EEG data for alcoholism detection. Utilization of perceptron-back propagation (MLP-BP) and probabilistic neural network (PNN) for alcoholic identification [19] suggest that though in a normal case the gamma band of EEG signal lies below 30 Hz, it can generate frequency between 30–50 Hz in case of an alcoholics. Support vector machine and neural networks were applied in [9] for alcoholism detection along with principal component analysis (PCA) for feature extraction. Recurrent neural network (RNN) is also reported to be used in EEG classification in numerous research [20, 21]. An automated diagnosis of alcoholics using numerous correlation function and support vector machine (SVM) was performed by [2]. The correlation functions identified the relation between different parts of the brain and if it changes in alcoholic condition. They have found that certain parts of the brain communicate in normal decision-making process and in an alcoholic condition the communication reduces significantly.

In this research, we have considered an open source dataset [3] of EEG signals. In a previous study by Ruslan Klymentiev¹ on the same dataset showed that among all the electrodes the most significant correlation is seen between FPZ & FP1 and FP1 & FP2 around 90% [3]. Rather than finding the correlation, we have applied ML algorithms directly on the response value received from these electrodes and see if they are significant enough to distinguish between alcoholics and healthy control.

3 Materials and Methods

3.1 Overall System

In this experiment we have applied ML methods for prediction of alcoholism based on EEG signals. To classify alcoholics from healthy control and com-

¹ <https://www.kaggle.com/ruslankl/eeg-data-analysis>. Accessed on March 23th, 2020.

pare the performance of instance based learning and neural networks, we have designed a step by step system. A snapshot of the overall methodology of the system is provided in Fig. 1.

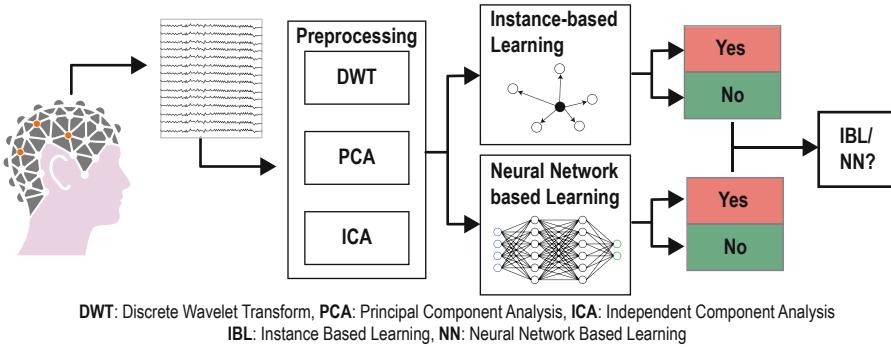


Fig. 1. Block diagram of the implemented pipeline showing different steps and execution sequence.

3.2 Dataset Description

For our experiment, we considered EEG dataset which was gathered to examine EEG correlates of genetic predisposition to alcoholism [3]. The characteristics of this dataset is multivariate, time series which has attributes with categorical, integer and real valued properties. The EEG was collected while the two groups of subjects were shown a set pictures from the 1980s Snodgrass and Vanderwart picture set [3]. The EEG were recorded from 64 electrode placement (10–20 standard EEG placement) on each subject. Each subject was exposed to either a single stimulus (S1) or to two stimuli (S1 and S2) and were asked to identify either a matched condition where S1 and S2 were identical or in a non-matched condition where S1 and S2 are different. The EEG recorded was sampled at 256 Hz (3.9 ms epoch) per second. The original dataset contains EEG of 20 subjects where each subject completed 120 trials for each stimulus. For our experiment we have randomly selected 5 alcoholics and 5 healthy controls among them. Then the response values of FPz, FP1 and FP2 are collected for all 120 trials of each given stimulus. The training and test data were prepared using stratified 10-fold cross validation.

3.3 Data Preprocessing

Data preprocessing is an important concepts in EEG data. In our case, EEG data shows high variance which required noise cleaning to get closer values. The high level idea is to enhance the likelihood for producing a cut above result. Our objective of research is to compare instance-based classifiers with NN to demonstrate variance of result in respect to alcoholism diagnosis. Hence cleaner data

can potentially aid in experiment for making informed decision on model accuracy. Our inclusion criteria for choosing DWT, PCA and ICA were: first reviewing the evidence of efficacy from literature and the efficiency of their underlying mathematical model for computation. There are particular distribution based on different data source which are respectively Gaussian with different kurtosis. EEG data has a distribution of super-Gaussian which needs application of transformation that pass distribution of non-linearity to calculate entropy. ICA and PCA are quite similar based on their functionality. PCA is applied to the training data set to indicate transformation matrix which are used for measuring the final feature. For example, if transformation matrix is m with the dimension of $K \times N$, the outcome y will be: $y = m^T x$; where x is original vector by orthogonal basis where PCA help in feature reduction for our EEG dataset [7]. DWT offers a compressed approximation of the data that can be retained in a reduced representation of the original data. DWT can be also used for noise reduction by filtering out any particular order coefficients using a threshold.

DWT. Wavelet Transformation [18] is used for decomposition and summarizing a time-domain signal into a multidimensional representation comprised of a set of basis functions called wavelets. The wavelets are generated by scaling and shifting a mother wavelet. If transformation includes a discrete set of wavelets which are orthogonal to its translation and scaling, they are known as DWT. The DWT coefficient ($A\varphi[i, P]$) of a signal $x[n]$ is defined as:

$$A\varphi[i, P] = \frac{1}{\sqrt{k}} \sum_n (x[n] \Pi_{i,P}[n]) \tag{1}$$

Here, K is the number of samples and A is a wavelet function.

PCA. PCA [1] converts a set of observations of correlated variables into a set of values using an orthogonal transformation. The newly generated values are linearly uncorrelated and called principal components (PCs). Each PC must be orthogonal to its preceding components. If we have a matrix T of $p \times q$ then covariance matrix can be calculated as:

$$C_T = \frac{1}{n-1} (T - \bar{T})(T - \bar{T})^T, T^T \text{ is the transpose matrix of } T \tag{2}$$

From the covariance matrix C the eigenvalues and orthogonal eigenvector matrix P is calculated. Because the covariance matrix is positively semi-defined and symmetric in nature, the diagonal matrix Π is defined as $C_T = P \Pi P^T$. The eigenvalues are contained in Π successively corresponds to the values contained in eigenvector P .

ICA. ICA [7] is a popular blind source separation method that finds a linear representation of non-Gaussian data in a signal statistically independent components (ICs). ICA can identify the original signal from noise at least to a certain

level if some information about the origin of the signal is known [7]. To understand the concept of ICA lets consider an observed signal $x_i(t)$ represents a mixture of n signals. $x_i(t)$ can be modeled as:

$$x_i(t) = \sum_{j=1}^m \delta_{ij} S_j(t) \quad (3)$$

where, δ_{ij} is a constant parameter called mixing matrix $s_j(t)$ represents an IC at time point t . s is the source signal to be separated from its mixed component δ_{ij} . Denoting the elements δ_{ij} as T , it's inverse matrix W is calculated to obtained the ICs as $s = Wx$. The ICs generated must have non-Gaussian distribution and their number is equal to the number of observed sources.

3.4 Machine Learning Algorithms

Two types of ML models are used in this study: instance-based (IBL) and neural networks (NN)-based. IBL algorithms do not create any learning model before the actual classification process, thus, without separating the training and testing phases and creating a model local to certain test tuple or instance. This study applies two IBL models: K-Nearest Neighbor (KNN) and Learning Vector Quantization (LVQ). The reason to choose KNN and LVQ is that they are most commonly used IBLs and have been reportedly used in time series data classification. On the other hand, NN is brain-inspired, where in each layer of the network, the neurons learn from a training dataset. Later this model predicts classes for given queries. Here three well known NNs will be applied Recurrent Neural Network (RNN), Bidirectional Long-Short Term Memory (B-LSTM) and Convolutional Neural Network (CNN).

Instance-Based Methods

KNN. KNN is one of the popular ML models. KNN does not necessarily create a classifier model from the input space. When a query is fed into the classifier, KNN algorithm chooses its k nearest neighbors by calculating the distance between the query and other instances. From the number of methods for distance calculation, commonly used distance metric calculation method is the Euclidean distance. Euclidean distance is the rooted sum of squared distance between two values of the same attribute. For two instances X and Y with I attribute the Euclidean distance between them can be calculated as:

$$E = \sqrt{\sum_{i=1}^n (X_i - Y_i)^2} \quad (4)$$

After selecting the k nearest neighbor the class label of the instance with which the query has the lowest distance will be assigned to the query.

LVQ. LVQ allows to determine exactly how many instances from the training set is needed to learn generating a more optimized classification model [8]. This particular set of instances is called the “window”. The window is determined around the mid-plane of two variables m_x and m_y . m_x and m_y are two nearest neighbors of the query q such that if m_x and q belong to the same class, m_y and q will have different class label and vice-versa. The relation between m_x , m_y and q can be defined with the following equation:

$$m_x(i+1) = m_x(i) - \alpha(i)[q_i - m_x(i)] \quad (5)$$

$$m_y(i+1) = m_y(i) - \alpha(i)[q_i - m_y(i)] \quad (6)$$

Here, $\alpha(i)$ is individual learning rate factor. The value of m_x and m_y is updated at each step i until the closest instance m_x is found. Finally, the class label of m_x is assigned to the query q . In this experiment we have applied LVQ3 as it is more robust comparing to LVQ1 and LVQ2. It provides both binary and multi-modal classification.

Neural Network Based Methods

RNN. RNN is a modified feed-forward neural network which has an internal memory that contains information about previously learned data. At a certain hidden layer it makes decision from the current input and outputs from previous layer. Traditionally, the state of a hidden layer neuron of RNN is computed as:

$$A_i = A(A_{i-1}, x_t) = W_A A_{i-1} + W_x x_i + b \quad (7)$$

Here, W represents a weight parameter. At each hidden step A_t , the output is calculated using an activation function that is applied on input x_i of the current layer and output A_{i-1} of the previous layer. In this experiment we have used the modified version of RNN called long-short term memory (LSTM) [5] that can easily model time sequenced data such as EEG. RNN shows two long term dependency problem, the vanishing gradient problem and the exploding gradient problem. These problems can be handled by using LSTM. It uses designated hidden states called cell that stores information for long period of time so that particular information is accessible to both immediate subsequent steps and later nodes. It's special gates can control removing or adding information to a cell state. It has three specialized gates called the forget gate (F_i), input gate (I_i), and output gate (O_i). Each gate produces an output using similar equation to a RNN hidden gate. The final Output of an LSTM cell with these gates can be defined by:

$$A_i = O_i \otimes \tanh(L_i) \quad (8)$$

Here, L_i represents recurrent state of the LSTM node and has following form

$$L_t = L_{i-1} \otimes F_i \oplus \tilde{L}_i \otimes I_t \quad (9)$$

Our designed LSTM network The second layer uses activation function sigmoid that returns a number between 0 and 1 depending on the cell state.

B-LSTM. Bidirectional LSTM is an advanced LSTM that learns not only from the previous layer but also from the future elements. Therefore, instead of one recurrent network it trains two, respectively for previous and future outcomes. The input sequence is fed to one network in normal time order and the in-reversal time order for the other one. Both outputs are concatenated or summed at each time step to generate the current state. B-LSTM might use similar activation functions as LSTM.

CNN. CNN uses layers of convolution that convolve a filter also called window over an input dataset and generates a feature map where some activation function is applied. In convolution network, Convolution layer calculates the output of neurons, connected to local regions in the input by applying a dot product between their weights and values of the input volume in the local region. It is followed by a pooling layer that uses some aggregation function to create a pooled map along the spatial dimensions to reducing the size of the connected layer. Finally, the fully connected layer computes the classification score. In CNN multiple convolution and pooling layer is applied alternatively to create a more accurate classification model or network. In our study, we have applied 1-dimensional convolution layer where each instance acts as a input vector. A pooling layer of pooling size 2 is applied.

3.5 Evaluation Metrics

The classification result is evaluated using confusion matrix which represents the number of correct and wrong prediction for each ML algorithm. Based on the confusion matrix different performance metrics are calculated (Table 1).

Table 1. Performance evaluation metrics

| Measurement | Formula |
|---------------|-------------------------------------|
| Accuracy | $\frac{TP+TN}{TP+TN+FP+FN}$ |
| F-measure | $\frac{2TP}{2TP+FP+FN}$ |
| G-mean | $\frac{TP}{\sqrt{(TP+FP)*(TP+FN)}}$ |
| Sensitivity | $\frac{TP}{TP+FN}$ |
| Specificity | $\frac{TN}{TN+FP}$ |
| Type I Error | $1 - \frac{TN}{TN+FP}$ |
| Type II Error | $1 - \frac{TP}{TP+FN}$ |

Table 2. Comparison of Instance-based and Neural Network Models

| Type | Model | Acc | ROC | F-M | G-M | Sens | Spec | T1E | T2E |
|------|---------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|
| IBL | KNN | 0.731 | 0.731 | 0.733 | 0.733 | 0.737 | 0.725 | 0.242 | 0.274 |
| | LVQ | 0.728 | 0.728 | 0.732 | 0.732 | 0.742 | 0.714 | 0.257 | 0.285 |
| NN | LSTM | 0.898 | 0.898 | 0.899 | 0.899 | 0.908 | 0.873 | 0.091 | 0.112 |
| | B_LSTM | 0.950 | 0.950 | 0.950 | 0.950 | 0.945 | 0.955 | 0.054 | 0.044 |
| | CNN | 0.850 | 0.850 | 0.855 | 0.855 | 0.882 | 0.819 | 0.117 | 0.180 |

Legend: Acc.: Accuracy; ROC: Receiver Operating Characteristics; F-M: F-Measure; G-M: Geometric Mean; Sens.: Sensitivity; Spec.: Specificity; T[1/2]E: Type [I/II] error; bold faced values denote the best performance.

4 Results and Discussion

The ML methods as well as the preprocessing techniques were implemented using the scikit library with Python. The evaluation metrics were measured separately for raw dataset and each preprocessing technique. The training and test dataset were prepared using stratified 10-fold cross validation.

4.1 Overall Comparison of the Algorithms

The performance of the considered algorithms against different evaluation metrics is shown in Table 2. We see that, in case of IBL, KNN achieved an accuracy of 73% and LVQ’s accuracy was only 72%. On the other hand, in case of NN, LSTM, B-LSTM and CNN received accuracy of 89%, 95% and 85%, respectively. In terms of accuracy the NNs offer more accurate classification result than IBLs. Though both the IBL algorithms have moderate accuracy rate, they lag behind the NNs notably. In addition, if we observe the other performance metrics, LSTM and B-LSTM has the lowest error rate (both type I and type II error), around 9% and 5% of type I, and 11% and 4% of type II for LSTM and B-LSTM consecutively. The other NN algorithm CNN also has a low error rate of only 11% for type I and 18% for type II. However, both IBLs has higher error rate compared to all NNs applied. 24% of type I and 27% of type II error received for KNN and 25% of type I and 28% of type II error received for LVQ. The NNs shows better performance for other evaluation metrics as well (Table 2). Among the five algorithms from both IBL and NN, we can see that the best performance is received from B-LSTM which achieved an accuracy around 95% and highest result in other performance metrics. Between the two IBL methods the highest accuracy is obtained by KNN and it also shows better result in other metrics against LVQ.

4.2 Effect of Preprocessing on Model Performance

In this section, we take into account the best model from two type of models and discuss their performance after applying numerous data preprocessing techniques

and also for the raw data. KNN is observed to be the best IBL classifier and B-LSTM is observed to be the best NN classifier (Table 3). In case of KNN the highest accuracy (88%) is received when the dataset is transformed using ICA. For DWT, PCA transformed data and raw data, accuracies obtained respectively are around 72%, 73%, 74%. As for B-LSTM, the highest accuracy achieved is for PCA transformed data (95%). DWT (accuracy 88%) and ICA (82%) did not necessarily improve the classification performance of B-LSTM, as the accuracy they achieved are lower than the accuracy received from B-LSTM when applied to raw data. The performance is similar for other evaluation metrics also.

B-LSTM outperforms KNN in case of raw data and other preprocessed data except for ICA (Table 3). When the dataset is preprocessed with ICA, KNN receives an accuracy of nearly 89% which is quite higher than B-LSTM who receives an accuracy of 82%. Based on the discussion above, we find that neural networks show better classification compared to instance based learning for alcoholism prediction with time series EEG data and B-LSTM is observed to be the best classifier in this experiment. However, there is one drawback observed of NNs which is each of them requires higher run-time than IBL algorithms.

4.3 Discussion

In this study we applied two groups of ML algorithms (KNN & LVQ for IBL and RNN, B-LSTM & CNN for NN) on EEG data to distinguish alcoholics from healthy controls. In addition, we compared the performance of these two groups using different performance metrics. There have been a number of previous studies analyzing EEG signals with ML for alcoholism detection. In this section we compare our study with some previous works regarding EEG analyzing with ML. In a study of automatic diagnosis of alcohol abuse [13] has shown significant difference between alcoholics and healthy controls in their EEG specially

Table 3. Effect of transformation on the performance of IBL and NN models

| Model | PPM | Acc | AUC | F-M | G-M | Sens | Spec | T1E | T2E |
|-----------------------|------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|
| IBL best model, KNN | DWT | 0.720 | 0.720 | 0.721 | 0.721 | 0.724 | 0.715 | 0.275 | 0.284 |
| | PCA | 0.731 | 0.731 | 0.733 | 0.733 | 0.737 | 0.725 | 0.242 | 0.274 |
| | ICA | 0.889 | 0.889 | 0.888 | 0.888 | 0.888 | 0.890 | 0.111 | 0.109 |
| | Raw | 0.747 | 0.747 | 0.745 | 0.745 | 0.739 | 0.754 | 0.260 | 0.245 |
| NN best model, B-LSTM | DWT | 0.885 | 0.885 | 0.887 | 0.887 | 0.898 | 0.872 | 0.101 | 0.127 |
| | PCA | 0.950 | 0.950 | 0.950 | 0.950 | 0.945 | 0.955 | 0.054 | 0.044 |
| | ICA | 0.827 | 0.827 | 0.828 | 0.828 | 0.835 | 0.818 | 0.164 | 0.181 |
| | Raw | 0.920 | 0.920 | 0.919 | 0.919 | 0.910 | 0.930 | 0.089 | 0.069 |

Legend: PPM: Preprocessing Method; Acc.: Accuracy; ROC: Receiver Operating characteristics; F-M: F-Measure; G-M: Geometric Mean; Sens.: Sensitivity; Spec.: Specificity; T[1/2]E: Type [I/II] error; Bold are best performance.

in the left hemisphere. Besides discriminating EEG of alcoholics and healthy controls, they have also differentiated the EEG of alcoholics and alcohol abuser though there was no seemingly significant difference between the alcoholics and alcohol abusers. However, EEG of alcoholics and healthy controls have high difference in the delta and theta band. They have received an accuracy of 96% using SVM. In another study [15], wavelet transformation methods is applied to find non-linear correlation called correntropy in EEG signals of alcoholics and normal. The correlation is then used with Squared SVM for classification which received an accuracy of 97%. However, their study does not provide any detail discussion on if EEG of any certain part of the brain carries any differentiating features between the alcoholics and the controls. In our study, We have considered response values of electrodes FPZ, FP1 and FP2 which collects neural activity in the prefrontal cortex of the brain. The prefrontal cortex is the part of the cortical region responsible for decision making as well as reasoning [4] based on past events. Therefore, we have considered the fact that neural response of alcoholics and the controls can have significant difference in their prefrontal cortex. In case of accuracy, both the papers [15] and [13] have marginally higher performance than ours where our best classifier B-LSTM have received an accuracy of 95%. However, the aim of this study was not only classify the EEG data, also compare the performance of IBL and NN algorithms. We have applied different preprocessing methods to both groups of methods and found that NN outperforms IBL in classifying non-linear EEG data, except when ICA is used.

5 Conclusion

Alcoholism is a psychological phenotype harmful to an individual as well as to society. The negative physical effects of its' can be transmitted genetically to the offspring. Therefore, identification of alcohol abusers from healthy people is becoming an important research topic to data scientists. Our experiment successfully implemented ML algorithms for classifying EEG data of alcoholics and healthy control. Neural networks outperformed instance based algorithms, however if the time-series EEG data is converted to linear data using ICA, the instance based networks significantly improve in classifying EEG of alcoholics, even outperform the neural networks. B-LSTM is proven to be the best classifier in this experiment receiving accuracy of at most 95%. Our experiment of comparing two different classifiers' performance with EEG signal suggest the best possible method for Alcoholism automated detection and therefore, reduce the diagnosis error. In the future we want to use the response values of all the patients that participated in the EEG data collection and see if the algorithms show scalability.

References

1. Abdi, H., Williams, L.J.: Principal component analysis. *Wiley Interdiscip. Rev. Comput. Stat.* **2**(4), 433–459 (2010)
2. Acharya, U.R., Sree, S.V., Chattopadhyay, S., Suri, J.S.: Automated diagnosis of normal and alcoholic EEG signals. *Int. J. Neural Syst.* **22**(03), 1250011 (2012)
3. Begleiter, H.: EEG database data set. Neurodynamics Laboratory, State University of New York Health Center Brooklyn, New York (1995). <https://archive.ics.uci.edu/ml/datasets/EEG+Database>
4. Fuster, J.: *The Prefrontal Cortex*. Academic Press, Cambridge (2015)
5. Hochreiter, S., Schmidhuber, J.: Long short-term memory. *Neural Comput.* **9**(8), 1735–1780 (1997)
6. Hommer, D.W., et al.: Evidence for a gender-related effect of alcoholism on brain volumes. *Am. J. Psychiatry* **158**(2), 198–204 (2001)
7. Hyvärinen, A., Oja, E.: Independent component analysis: algorithms and applications. *Neural Netw.* **13**(4–5), 411–430 (2000)
8. Kohonen, T.: Learning vector quantization. In: Kohonen, T. (ed.) *Self-Organizing Maps*. Springer Series in Information Sciences, vol. 30, pp. 175–189. Springer, Heidelberg (1995). https://doi.org/10.1007/978-3-642-97610-0_6
9. Kousarrizi, M.R.N., Ghanbari, A.A., Gharaviri, A., Teshnehlab, M., Aliyari, M.: Classification of alcoholics and non-alcoholics via EEG using SVM and neural networks. In: *Proceedings of ICBBE*, pp. 1–4. IEEE (2009)
10. Li, M., et al.: Emotion recognition from multichannel EEG signals using k-nearest neighbor classification. *Technol. Health Care* **26**(S1), 509–519 (2018)
11. Mahmud, M., Hawellek, D., Bertoldo, A.: EEG based brain-machine interface for navigation of robotic device. In: *Proceedings of BioRob*, pp. 168–172. IEEE (2010)
12. Mahmud, M., Kaiser, M.S., Hussain, A., Vassanelli, S.: Applications of deep learning and reinforcement learning to biological data. *IEEE Trans. Neural Netw. Learn. Syst.* **29**(6), 2063–2079 (2018)
13. Mumtaz, W., et al.: Automatic diagnosis of alcohol use disorder using EEG features. *Knowl. Based Syst.* **105**, 48–59 (2016)
14. Oh, S.L., et al.: A deep learning approach for Parkinson’s disease diagnosis from EEG signals. *Neural Comput. Appl.* **32**, 1–7 (2018)
15. Patidar, S., Pachori, R.B., Upadhyay, A., Acharya, U.R.: An integrated alcoholic index using tunable-Q wavelet transform based features extracted from EEG signals for diagnosis of alcoholism. *Appl. Soft Comput.* **50**, 71–78 (2017)
16. Santa Maria Shithil, T.K.S., Sharma, T.: A dynamic data placement policy for heterogeneous hadoop cluster. 2017 4th International Conference on Advances in Electrical Engineering (ICAEE) (2017)
17. Shanir, P.M., Khan, K.A., Khan, Y.U., Farooq, O., Adeli, H.: Automatic seizure detection based on morphological features using one-dimensional local binary pattern on long-term EEG. *Clin. EEG Neurosci.* **49**(5), 351–362 (2018)
18. Shensa, M.J., et al.: The discrete wavelet transform: wedding the a Trouse and Mallat algorithms. *IEEE Trans. Sig. Process* **40**(10), 2464–2482 (1992)
19. Sriraam, N., et al.: EEG based detection of alcoholics using spectral entropy with neural network classifiers. In: *Proceedings of ICoBE*, pp. 89–93. IEEE (2012)
20. Stoev, H.: Brain disease detection from EEGs: comparing spiking and recurrent neural networks for non-stationary time series classification. Masters dissertation. Technological University Dublin (2020). <https://doi.org/10.21427/sv9j-t268>
21. de Zambotti, M., et al.: 0141 effect of evening alcohol intake on polysomnographic sleep in healthy adults. *Sleep* **41**, A55 (2018)