# Alcohol Use Disorders Automatic Detection based BCI Systems: A Novel EEG Classification based on Machine Learning and Optimization Algorithms

Said Abenna, Mohammed Nahid, Hamid Bouyghf Hassan II University, Faculty of Sciences and Technology, Casablanca, Morocco said.abenna@ etu.fstm.ac.ma, mohammed.nahid@ gmail.com, hamid.bouyghf@ gmail.com

Abstract—Alcohol is a serious toxic substance that alters brain function by interfering with neuron processes in the central nervous system, leading to mental and behavioural disorders. Alcoholism has serious pathological effects on the liver, immune system, brain, and heart. Alcoholism diagnosis is critical not only because of the disease's impact on individuals and society but also because of the financial cost to the national health system since people from it around manv suffer the world. Electroencephalogram (EEG) signals of normal and alcohol can be classified automatically to diagnose these illnesses, such that this work contains a simple and very fast prediction system. The method uses a bandpass filter to remove all unused signal frequencies, it can be showed in correlation matrices. Therefore, the use of machine learning (ML) algorithms in this work make it possible to quickly generate prediction models with better accuracy values, in addition to the improvement of the algorithmic parameters of classifier and bandpass filter using major optimizers like Genetic Algorithm (GA) and Harris Hawks Optimization (HHO), the elaborate model precision values have been increased from 63.9% to 99.95% without the extraction of EEG signal characteristics. To minimize the number of the electrodes and remained good accuracy values, this work uses the Extra-Trees (ET) algorithm, such as with only four electrodes the accuracy value remains higher of 99%. A comparison with other techniques was performed aiming to validate our approach, and it shows great efficiency, simplicity, and instantly.

*Index Terms*—Brain-Computer Interface (BCI), Electroencephalogram (EEG), Data analysis, Signal processing, Feature selection, Machine learning, Optimization.

#### I. INTRODUCTION

Alcoholism or poor alcohol consumption is a common neurological disorder that affects about 10% of the world's population [56], the most common negative health effects of excessive alcohol consumption are cardiomyopathy, stroke, high blood pressure, cirrhosis, and increased cancer risk. Many models mark a person as having a change in alcohol dependence is thought to suffer from alcoholism throughout a continuous alcohol disorder, on the other hand, Churchill and Farrell [20] have also shown that there is a link between alcohol and depression [56], alcoholism also undermines the social life of people and their people and increases the cost of social investment in health care. In this pessimistic situation, male alcohol accounts for 7.6% of all deaths among the world's deaths [51, 56], it has a variables variety that affects the central and peripheral nervous systems, affecting their function and leading to the recoverability of normal brain activity [8, 56]. Moreover, alcoholism is considered a neurological disorder characterized by excessive alcohol consumption that does not control alcohol consumption [24, 45], where the alcoholism falls into three categories, light, moderate, and severe depending on the pattern and quantity of alcoholism, such as the light or moderate drinking can have a health benefits variety, including reducing the risk of diabetes and heart disease, while excessive drinking can lead to health problems variety, including stroke, liver disease, cardiovascular disease, cancer, and cirrhosis. According to the world health organization (WHO) [45, 65], when alcohol consumption is severe, the mortality rate has risen to three million, accounting for 5% of the global burden of disease and about two billion people drink alcohol, of whom 81.7 million are severely poisoned [12, 58, 59], manual screening of EEG signals is difficult and often produces inaccurate results [12, 49]. The brain-computer interface (BCI) converts the brain's electrical activity into meaningful commands using soft computing technology to identify the EEG signals tasks [1, 50, 64] and to control a variety of applications and assistive devices, such as robotic arms and electric wheelchairs for helping many of people with disabilities, this EEG-signals are essentially nonstationary varying in mental state and brain awakening level [1, 34, 64]. In addition, EEG-brain records of different human

Submission Date: 25 April 2021

Author's current affiliations:

S. Abenna is with the University Hassan II, Faculty off Sciences and Technology, Mohammedia, Morocco (e-mail: said.abenna@ etu.fstm.ac.ma).

M. Nahid is with the University Hassan II, Faculty off Sciences and Technology, Mohammedia, Morocco (e-mail: mohammed.nahid@ gmail.com).

H. Bouyghf is with the University Hassan II, Faculty off Sciences and Technology, Mohammedia, Morocco (e-mail: hamid.bouyghf@ gmail.com).

subjects can be used for personal identification using each individual's activities [46, 53, 64]. There are various neuroimaging techniques, namely, computed tomography (CT), magnetic resonance imaging (MRI), positron emission tomography (PET), an electroencephalogram (EEG) signals are used to diagnose brain abnormalities such as stroke, epilepsy, cancer, Alzheimer's disease, and alcoholism [14, 17, 45, 49]. Neurology assessment is more convenient when calculating the power spectrum of alcohol and non-alcoholic brain signals. However, the nonlinearity, time domain, or frequency of the EEG signals makes domain analysis difficult. Various timefrequency domain signal processing technologies based on Fourier transformation (FFT) allows for obtaining the different EEG waves such as Delta, Theta, Alpha, Beta, and Gamma [1, 12, 15]. To analyze medical data sets, machine learning (ML) methods were created. The medical diagnosis is obtained from the patient's medical history. The developed classifier may be used to diagnose fresh datasets with greater speed, accuracy, and reliability. Because of its high performance, ability to deal with missing data, ability to explain decisions, and knowledge transparency, the ML system is more useful in solving medical diagnosis problems [39, 40]. Different machine learning approaches can be used, and the literature reports on a wide range of techniques and how they can be used for classification. Classifiers based on Gaussian mixture models span from simple linear discriminant analysis (LDA) to nonlinear and highly sophisticated Gaussian mixture models [6, 55]. Several traditional methods, such as LDA [41], Long 2014, support vector machines (SVMs) [5, 68], [25, 66, 67], k-nearest neighbor (k-NN) [16], hidden Markov model [22], fuzzy systems [31, 38], etc. All classical classifiers have only one classifier in common. Ensemble ML classifiers, which mix many classical classifiers, have recently been presented as a way to increase the performance of a single classifier. The random forest (RF) introduced by Breiman is one of the most widely used ML algorithms, with applications in a wide range of fields [9, 16]. In present work uses a filter-based method that selects low-frequency signals between 63.572 and 64.572Hz [1], knowing that all noise sources are eliminated, making the classification stage easier, such as the machine learning algorithms that gives higher results because of the high speed of classification and prediction compared to other algorithms. Then the use of optimization algorithms like a genetic algorithm (GA) finds good parameters filling more stable and efficient [2]. As a result, the results in this work show the detected alcoholism fictitiously in a population with an accuracy value is more than 99% without using data decomposition algorithms. The rest of this article is organized as follows, section II present the related work for this paper, section III presents the dataset used, data processing, classification and optimization algorithms proposed for the prediction system, the results obtained and discussed are presented in section IV, while section V provides conclusions and an overview of future work.

## II. RELATED WORK

In the past, several studies have been conducted using EEG signals to identify alcoholism and non-alcoholic subjects.

Anuragi and Sisodia [12] have proposed a new machine learning framework based on Experience Wave Conversion (EWT) for classifying alcohol and normal subjects using EEG signals. Boundary detection methods are used to segment the Fourier spectrum of EEG signals represented in proportional space. Hilbert Huang transform (HHT) uses real-time amplitude (AI) and real-time frequency (IF) to check time and frequency information for a single region. AI and IF are used to form built-in pattern (IMF) functions. The conversion of stopgap waves (EWT) using Hilbert-Huang transform (HHT) extracts statistical characteristics from each built-in pattern function (IMF), such as mean, standard deviation, variance, distortion, Kurtosis, Shannon entropy, and record entropy. Extracted entities are evaluated by t-tests to find the most important entities. Important functional arrays are available for a classification algorithms variety that is listed as considering least square-SVM (LS-SVM), support vector machines (SVM), Nave-Bayes (NB), and k-nearest neighbours (K-NNs). Crossvalidation (LOOCV) is used to train and test used models to minimize the over-fitting possibility. Mehla et al. [45] uses a new method based on the Fourier theory, called the Fourier decomposition method (FDM), to automatically identify alcoholism using sensory signals (EEGs). The FDM method is used to decompose the EEG signal into many required orthodectonic elements, often referred to as Fourier inherent band functions (FIBF) and to obtain this by dividing the entire EEG signal band under analysis into equal frequency bands. Time-domain features such as Haworth parameters, kurtosis, quarter-to-quarter intervals, and average frequencies are extracted from FIFS. To reduce complexity, the Kruscal-Wallis (KW) statistical test uses the most important features. Tunable-Q wavelet transform (TQWT) features collected from EEG data were used by Patidar et al. [52] to present a new method for diagnosing alcoholism. The centred corr-entropy (CC) is extracted from the fourth decomposed detail sub-band using TQWT-based decomposition. For feature reduction, PCA is applied, and LS-SVM is used to identify normal and alcoholic EEG signals. A 10-fold cross-validation approach is used to ensure that classification performance is dependable. Rodrigues et al. [56] have announced a new classification of EEG Alcohol Signals using wave subcontract degradation (WPD) and Machine Learning Technology. The experiment uses the minimum value, high point, means, standard deviation, power value, absolute mean, and absolute mean as the characteristics to provide the classifier. These characteristics are combined with the ability to explore the ability to perform such characteristics to classify alcoholism. Fastening tasks are performed using SVM, optimal path forest (OPF), Naive Bayes, k-nearest-neighborhood (k-NN), and multi-layer-perceptron (MLP). Upadhyay et al. [64] propose a method for extracting EEG signals characteristics using continuous-wave conversion and suggests and validates two different brain states for identification: alcoholism and normality. The wave continuous conversion coefficient using four different base waves is calculated from the processed EEG signals. In addition, the statistical parameters of each basic wave are calculated, and the feature carrier is generated from EEG. Anuragi and Sisodia [10]



Fig. 1. Proposed general architecture for alcoholism dataset classification.

used a nonlinear parameter signal support vector computer technology in the characteristic time-frequency domain extracted from the EEG signals. The function uses continuous wave conversion extraction, and the adjusted Q-wave transformation (TQWT) is used for signal disassembly and to extract centralized core anthropology (CC) functions (which are subs slices of decomposition) and to look for small changes in nonlinear signals with delay time delays, much like signal autoassociation. These features are reduced by applying PCA and passed to an LSSVM for EEG signals classification of alcoholic and normal. Taran and Bajaj [63] propose an EEG rhythmbased feature for automatically identifying and analyzing alcohol EEG signals. Analysis by the built-in pattern function (IMF) represents that the calculated instantaneous frequency is used to isolate different frequency ranges called EEG rhythms. The IMF obtains this by applying the decomposition of empirical models to EEG signals. The variability and complexity of isolated EEG rhythms are measured by characteristics such as mean absolute deviation, quarter range, variance coefficient, entropy, and inverses. The P-value analysis of these features reveals that low-frequency rhythm (LF)-based characteristics have high statistical relevance for the identification and neurological interpretation of alcohol EEG data. The LF rhythm function serves as an input to the ML algorithm, an LS-SVM classifier used to classify normal and alcoholic EEG signals. Mumtaz et al. [48] use static EEG derivation as input data for recommended functional selection and classification methods. The goal is automatic classification and health control of alcohol use disorder (AUD) patients. The validation of the proposed method included actual EEG data obtained from health controls that matched AUD patients. Functions extracted from rest EEG, such as synchronization likelihood (SL), are calculated to contain 19 scalp positions and create 513 functions. In addition, this feature is commanded to select the most discrete features according to the standard, including the range-based functional selection method, the receiver operating characteristics (ROC). Therefore, when classifying AUD and healthy control patients, a small group of the most discrete characteristics was found and used. Mumtaz

et al. [47] propose an ML approach that ranks alcoholics and health controllers (1, 2) with health control, alcoholics, and alcoholics (1). Suggested ML methods include extracting QEEG characteristics, selecting the most relevant features, and classifying them as relevant study participants. Acharya et al. [7] present a nonlinear characteristics-based computer-aided diagnosis (CAD) technique for automatically recognizing normal and alcoholic EEG signals. First, it is extracted and used to train nonlinear features such as SVM classifiers (1st, 2nd, and 3rd polyhedral and radial substations) (RBF), such as approximate entropy (ApEn), maximum Leaponov index (LLE), sample entropy (SampEn), and four other higher-proof of stake functions (POS). Rachman et al. [54] recommends the use of a stand-independent component analysis (ICA) algorithm as a functional extraction method for the stationary wavelet transform (SWT) of notation and divides it into two categories, namely alcoholism and normality, using probabilistic neural network (PNN).

#### III. MATERIALS AND METHODS

#### A. Dataset



Fig. 2. Electrodes position for alcoholism detection dataset.

The dataset used in the study (KDD dataset) of EEG signal acquired during tests on subjects with a genetic predisposition

to alcoholism. The cognitive decline surveyed subjects who were not able to cool the alcohol, where all subjects were divided into two groups: control and alcohol, for the alcohol group, consisted of 77 male alcoholics with an average age of 35.83, with ages ranging from 22.3 to 49.8 years, on the other hand, the control group consisted of 48 subjects who had no specific family history of drugs or alcoholism or any history of mental or neurological disorders [56]. According to the international system 10-20, there are 61 non-invasive channels of the international electric cap for signal acquisition, where each electrode impedance is always below 5.0k $\Omega$ , and the subject is punished with a nasal electrode. The EEG signal is sampled at 256Hz during 190 milliseconds of the postulation baseline, 1440 milliseconds after each stimulus. The signal is a band that is transmitted between 0.02 and 50Hz using an EPA-2 amplifier. Stimulation was done under three conditions using the Snodgrass and Vanderwat acquisition datasets [56]:

 Condition 1: A single stimulus package (S1) was proposed to the subjects.

• Condition 2 (matching condition): Stimulus 2 (S2) is repeated as S1.

• Condition 3 (nonmatching condition): S2 is followed by another image of S1 in its meaning category.

Fig. 2 illustrates all 61 electrodes used to acquire EEG signals during the alcoholism detection experiment, we notice that all the electrodes are distributed over the entire scalp surface, this improves the signals acquired quality.

#### B. Data processing

This study used EEG data from all 122 subjects in the alcoholism dataset. The acquisition time was 11056 seconds and the sampling rate was 256 Hz. The compressed data size is only 40 seconds (10240 samples), this is guaranteed that the signal will sample one second of data every 276 seconds of the margin. The signal is applied to the gamma bandwidth (all frequencies from 40 to 80 Hz) with the filter bandwidth, signal and labels fs = 256Hz. shuffle (random-state = 123) and splitting the data by 80% for training, the remaining 20% for testing. The label shows the various EEG signal tasks obtained in the alcoholism dataset as tracking:

- Task 'A': The subject is an alcoholism
- Task 'C': The subject is a control (non-alcoholism).

1) Correlation matrix: The electrode correlation coefficient is a statistical measure of the linear correlation between two electrode signals.

$$corr_{xy} = \frac{\sum(X-\bar{X})(Y-\bar{Y})}{\left(\sqrt{\sum_{i=1}^{n}(X-\bar{X})}\right)\left(\sqrt{\sum_{i=1}^{n}(Y-\bar{Y})}\right)}$$
(1)

The correlation coefficient is expressed as 1, and n is the sample size, X,  $\overline{X}$ , Y, and  $\overline{Y}$  are observations and averages for the two variables, respectively. Describes the degree of linear correlation cor between two variables. Correlation coefficients range from 1 to 1.  $corr_{xy} > 0$  indicates that the two variables are positively related. If  $corr_{xy} < 0$  indicates a negative correlation between two variables, we can use the corr() method learned in scikit-learn to calculate the standard correlation coefficient

between each pair of channels or calculate the correlation between each property and classification label and select the most relevant primary channels data.

2) Band-pass filter: In this proposal, we are interested in applying filters to eliminate the simplest preprocessing techniques for undesirable frequencies. To do this, we chose Butterworth filters, which are widely used in many of BCI applications, because they produce uniform pitch bands. This filter, also known as the maximum plane filter, is widely used in different signal processing applications due to its functionality [1, 57]. The main advantage of Butterworth filters is that they respond by flat size in the band, but have the cost of a wide transition band. For high-speed four-wire centre conversion, we can apply a Butterworth bandpass filter in the frequency domain that represents H(u; v) in revenue.

$$H(u,v) = 1 - \frac{1}{1 + \left[\frac{D(u,v)W}{D(u,v)^2 - D_0^2}\right]^{2n}}$$
(2)

D(u; v) is the distance between the (u; v) point and the frequency rectangle center. W is the bandwidth  $D_0$  is the band center, n is the filter order [57]. For a N filter, the four-power function of the function equation size is: |D(3)

$$(j\omega)|^2 = 1 + \varepsilon \omega_n^{2N} \tag{2}$$

 $\omega_n$  represents the normalization frequency,  $\varepsilon$  represents the route band attenuation. We also do zero-phase digital filtering to avoid work parts. The algorithm returns a filtered signal

| Algorithm 1: Extra tree algorithm            |
|--|
| - Initialize the data D;                     |
| if $ D  < n_0$ or all characteristics remain |

constant in D or the output variable is constant in D then

- Add a new leaf node to the tree *Tree*;
- Save the example list D in Tree;
- Save the D class label distribution in Tree;
- Return the leaf node, where the *tree* is
- labeled by class frequencies in D;

## end

\_

- else
  - Using the split test s, divide D into two subsets:  $D_0$  and  $D_1$ ;
  - From these subsets, construct  $Tree_0$  as an extra tree  $(D_0)$  and  $Tree_1$  as an extra tree  $(D_1)$ .;
  - Create a node for decision-making. Attach  $Tree_0$  and  $Tree_1$  as left and right subtrees of this decision node using the split s.;

end

## return Tree;

indicating that an event is returned immediately after the original signal occurs [57].

#### C. Channels selection

Various ML algorithms enable determining the relevance

degree of each electrode for the classification stage, and channel selection are extremely crucial to reduce the number of acquiring electrodes without affecting the accuracy of the system [3, 4]. The Extra tree (ET) is used in this study to provide gradual learning by expanding the leaf nodes without recreating the entire tree. As a result, the classifier may be retrained using online streaming data [3]. Algo. 1 pseudocoding demonstrates the complete additional tree construction method. If the number of examples |D| is fewer than  $n_0$ , or if all candidate attributes in the example list D are constant, or if the output variable in D is constant, a new leaf node Tree is generated, containing the example list D and the class label distribution. A split test s is installed according to the split method in RF [28] when a leaf node is converted into a decision node, and two new descendant leaves  $Tree_0$  and  $Tree_1$  are generated. The old leaf node's example list is divided and populated into two new descendant leaves. The class label distribution on the new descendant leaves is also recalculated. The entire forest-building process is fairly quick since we only have to deal with a limited number of samples at a time [3]. For the feature selection based on Eq. (4), we can easily calculate the degrees of importance of each feature, so the normalized values are determined by Eq. (5).

$$FI[i] = \sum_{j \text{ in } F} nj * \sum_{k=1}^{Tree} f_k * (1 - f_k)$$
(4)

$$NFI[i] = \frac{FI[i]}{\sum_{k \text{ in } F} FI[k]}$$
(5)

## D. Classification

1) Gradient Boosting Model (GBM): Friedman's return mission was enlarged in 2011 when he introduced a gradient boosting system (GBM). The objective of GBM is to create an additive model that minimizes functionality loss [3, 4, 29]. As a result, Chen and Guestrin [18]'s extreme gradient booster (XGB) is an effective collection learning technique based on gradient boosted regression trees (GBRT) [35, 43]. The main concept behind cooperative learning is to generate a large number of underperforming elementary school students. GBRT in Algo. 2 detects error disparities between each weak student and the last regression tree using pulse techniques. The ultimate projected value for a particular input is the sum of all the results in the regression tree. The most essential aspect of XGBoost's success is its scalability across all circumstances. On a single machine, the system is 10 times quicker than existing popular methods, and it is scalable to billions of instances in distributed or memory-limited environments. XGBoost's scalability is due to several key systems and algorithmic improvements. A unique tree learning technique for managing sparse data is one of these advances, as is a theoretically justified weighted quantile sketch procedure for handling instance weights in approximation tree learning Chen and Guestrin [18]. Parallel and distributed computing speed up learning, allowing for more rapid model exploration. XGBoost also takes use of out-of-core processing, allowing data scientists to process hundreds of millions of samples on a single computer. Finally, combining these approaches to create an end-to-end system that scales to even

## Algorithm 2: GBRT algorithm

- Initialize the samples and labels of data 
$$(x,y)$$
;

- Initialize 
$$M_0(x) = \underset{\alpha}{\operatorname{argmin}} \sum_{i=1}^N \Phi(y, \alpha)$$
, with

 $\Phi$  is the loss function;

for *i*=1 to *N* do  
for *i*=1 to *n*-estimator do  

$$-y'_{ik} = -\left[\frac{\partial \Phi(y_i, M_{k-1}(x_i))}{\partial M_{k-1}(x_i)}\right];$$
  
 $-S_{tk} = T - nodeTree(y'_{ik}, x_i)$ , T is  
number of leaf,  $S_{tk}$  is a region of the  
value predicted for each leaf  $t;$   
 $-\alpha_{lk} =$   
 $\arg\min \sum_{x_i \in S_{tk}} \Phi(y_i, M_{k-1}(x_i) + \alpha);$   
 $-\alpha_{lk} = M_k(x_i) = M_{k-1}(x_i) + \nu * \alpha_{tk} 1(x \in S_{lk});$   
end  
end  
return  $M_k;$ 

greater data with the least amount of cluster resources is even more intriguing Chen and Guestrin [18].

2) Random Forest (RF): Random forests are a set of decision trees built from a guided sample of the original training set with a different set of subsets. The getting process randomly creates many different training sets from the original data set, each of which is used to training an individual decision-tree tree. During training, each decision tree node uses functions to separate and optimally classify in a randomly selected set of functions [2, 36, 42]. While creating the forest, each tree of solutions grows to the maximum depth without sorting on the square label and freely voting. As a final decision, forests will accept the highest voting class. The randomness of both training kits and selected tasks provides a variety of tree solutions, leading to high reliability against overfitting [16, 36].

*k-Nearest Neighbor (k-NN):* Most ML algorithms are used in conjunction with the KNN algorithm. The KNN classifier assigns the most frequent category label points in a specific category space k to the nearest training set [2]. The Euclidean distance between two points, calculated by the KNN classifier, is  $|\rho^p - \rho^q|$ , dimensions *x* and *x* + 1 can represent integral size d and delay time  $T_d$ , and these distances are given by [30]:

$$\sigma_x^2 = \sum_{d=0}^{x-1} [\rho^p (t + dT_d) - \rho^q (t + dT_d)]^2$$
(6)

Build feature vectors when training during classification, these vectors are represented by assigned samples, calculate the distance from the new vector to all-composite vectors, and select the nearest K samples.

3) Multi-Layer Perceptron (MLP): Multilayer perceptrons, called MLP networks, are mainly used to create classifiers. The main problem of using MLP networks is to determine the initial size of the network, associated with the number of hidden layers and neurons [2, 60]. MLP is a very popular artificial neural network technology that can map an input to output. The input

layer, hidden layer, and output layer are the basic building blocks of an MLP model. Back-propagation learning technology is used to train the network.

### E. Parameters optimization

1) Algorithm Genetic (GA): Genetic Algorithm (GA) is a parallel storage search optimization method developed by John Holland [32, 44], simulation of natural and bioethics genetic mechanisms in 1975. The main idea is to randomly generate early populations, constantly updating them through genetic operations such as selection, intersection, and mutation, and ultimately to obtain better solutions [21, 44]. The selection method of genetic algorithm features: First, encodes the entity parameters and gives the population an initial size decision. Each person in the population corresponds to a possible solution. Each person's fitness value is then calculated based on the competency value function. Set the probability of human forks and mutations [2, 44, 61]. The next generation is using appropriate genetic strategies, such as cross-rate, mutation rate, and selection rate until population yield reaches certain targets or completes a preset number of iterations [19, 44].

2) Harris Hawks Optimization (HHO): HHO is a new optimization algorithm, and its capabilities have not been widely studied in practice. The HHO algorithm was inspired by predatory birds in nature, such as Harris hawks, for their cooperative behaviours and pursuit strategies, and gives several escape models for the victim's movement and jumping. To solve optimization problems, various equations modelled after the social intelligence of Harris hawks were developed. Twenty-nine unconstrained benchmark tasks were used to evaluate HHO's performance. The avoidance of exploitative, exploratory, and local optima of HHO was investigated using unimodal, multimodal, and composition tasks. It was capable of discovering remarkable solutions when compared to other well-known optimizers [2, 13, 33].

## IV. RESULTS AND DISCUSSION

The experiments were conducted on the 2.4 GHz desktop and 6 GB of RAM with four Intel®Core (TM) i5 CPUs and 64 bit/Windows 10 operating system. TABLE I SYMBOL AND DESCRIPTION Symbol Description

| Symbol | Description            |
|--------|------------------------|
| N-est  | n-estimators           |
| L-rate | learning-rate          |
| RS     | random-state           |
| NEEG   | Number of Electrodes   |
| AC     | Accuracy               |
| SE     | Sensitivity            |
| PR     | Precision              |
| F1     | F1-score               |
| Κ      | Kappa-score            |
| МС     | Matthews Correlation   |
| JS     | Jaccard-Score          |
| CN     | Complete Ness          |
| NM     | Normalized Mutual Info |
| Н      | Homogeneity            |
| FM     | Fowlkes Mallows        |
| EV     | Explained Variance     |
| VM     | V-Measure              |
| MI     | Mutual Info            |
| ZOL    | Zero One Loss          |
| Тс     | Classification time    |
| Тр     | Prediction time        |
| То     | Optimization time      |
| Sc     | Classification speed   |
| Sp     | Classification speed   |



Fig. 3. Example of (A,C) binary-class confusion matrix



Fig. 4. A part of EEG for each binary-class (alcoholic or normal) acquired using 61 electrodes

Special Issue on Smart Cities, Optimization and Modeling of Complex Systems – iJIST, ISSN : 2550-5114 Vol. 6 - No. 1 - February 2022



Fig. 5. Correlation matrix for alcoholism detection EEG signals using the best 22 electrodes



Fig. 6. Channels selection classification for alcoholism dataset before and after the application of band pass filter

#### A. Model evaluation

The classification performance of into-concentrated classes is the most common technique of comparison algorithms. Therefore, there are some standard metrics in the field of computer learning to evaluate and compare the performance of different methods [37, 62]. Precision (PR), sensitivity (SE), accuracy (AC), and F1-score (F1) are considered to evaluate the classification performance of individual sleep stages representing the true-positive, true-negative, false-positive, and false-negative numbers of TP, TN, FP, and FN, respectively (Fig. 3.a). Accuracy in Table I is the most commonly used measure of experience, and it does not distinguish between different types of correct labels:

$$AC = \frac{TP + TN}{TP + TN + FP + FN}$$
(7)

Instead, there are two measures that are clearly close to the display of different classifier classes:

$$PR = \frac{TP}{TP + FP} \tag{8}$$

$$SE = \frac{TP}{TP + FN} \tag{9}$$

$$SP = \frac{TN}{TN + FP} \tag{10}$$

Sensitivity and specificity are commonly used to measure the performance of biomedical and complex data in the classification process [37, 62]. In this paper, sensitivity and specificity define the accuracy of detecting focal length and non-higher signals, respectively. Another measure used to evaluate performance is the measurement of F1 [62], then write as follows:

$$F1 = \frac{2*PR*SE}{PR+SE} = \frac{2*TP}{2*TP+FP+FN}$$
(11)

Cohen's Kappa (kappa) and Matthew's correlation coefficients (MC) are values of 1and10 that represent random correlations, while large positive (negative) values represent good (low) predictive quality for a particular structural class:

$$MC = \frac{TP*TN - FP*FN}{\sqrt{(TP+FP)(TP+FN)(TN+FP)(TN+FN)}}$$
(12)

$$\begin{cases} P_e = \frac{(TP+FP)(TP+FN) + (TN+FP)(TN+FN)}{(TP+TN+FP+FN)^2} \\ k = \frac{AC - P_e}{1 - P_e} \end{cases}$$
(13)

 $P_e$  or the possibility of an unexpected hypothetical agreement. The prediction speed of a generated prediction model is an essential parameter that makes it possible to evaluate the performance of the generated models using different classifications according to the maximum number of predictions per second probably during real-time testing, where it can be calculated by:

$$S_p = \frac{test - data - size}{T_p} = \frac{100*(FP + FN)}{T_p*(100 - AC)}$$
(14)

The classification speed of a classifier is a parameter that makes it possible to determine the degree of ease of the prediction models during training, which makes it possible to improve the speed of optimization of the different parameters of the global systems towards extremum states, where it can be calculated by:

$$S_{c} = \frac{train - data - size}{T_{c}} = \frac{4*100*(FP + FN)}{5*T_{c}*(100 - AC)}$$
(15)

#### B. Data analysis results

Fig. 4 shows some EEG signals acquired for the different classification tasks before and after employing the bandpass filter, we notice that the signals in Fig. 4.a (Alcoholic) are more active compared to those in Fig. 4.b (Normal). Fig. 4.c shows that all signals are less noisy and that the number of the picks is very high compared to the normal state due to an increase in eyes' blink, which means the possibility of detecting the difference between the tasks from signals of the electrodes (CP1; PZ; O1; AF8; C1; F1; ... etc.) easily.

1) Correlation matrices analysis: Fig. 5 shows correlation matrices of the twenty best electrodes selected with and without employing the filter, such that shows the links degree between every two electrodes signals. From this figure, we notice that the correlation matrix is changed completely after the application of the filter, as the best electrodes are possessed at values between -0.5 and 0.5. In Fig. 5.a, the best electrodes are (CPZ and C4), on the other hand, in Fig. 5.b, the best electrodes are (F8, FP2, and F5) compared to the other electrodes.

2) Channels selection results: Fig. 6 shows an order of all electrodes contributions when the classification phase, such that without the use of the filter (Fig. 6.a), the degree of all importance electrodes is between 0.014 and 0.02, knowing that the best electrode is PO8 and the bad is AF1, after the use of the filter (Fig. 6.b), we notice a great variance in the contribution degree of all electrodes is between 0 and 0.13, such that the best electrode is P8 and the bad is FT7. Fig. 6 shows the results of channel selection before and after the application of optimized filter using the Extra Tree algorithm, such that before the application of the filter in Fig. 6.a, we note that all channels with degrees of importance between 0.014 and 0.019, with only that the channel PO8 has a degree of 0.025, which shows that it is

TABLE II FILTER PARAMETERS OPTIMIZATION USING HHO AND GA FOR (A, C) CLASSIFICATION

|           |           |           | Class  | ificatio | on res | ults  |
|-----------|-----------|-----------|--------|----------|--------|-------|
| Optimizer | fb        | fh        | AC (%) | k (%)    | ZOL    | To(s) |
| GA        | 41.469803 | 68.932422 | 63.92  | 17.63    | 739    | 1.28  |
| HHO       | 47.529807 | 50.649080 | 77.44  | 48.24    | 739    | 1.54  |
| HHO       | 43.850670 | 46.495947 | 78.37  | 50.22    | 443    | 2.90  |
| HHO       | 56.103037 | 58.533968 | 78.61  | 50.86    | 438    | 5.31  |
| HHO       | 77.105243 | 78.788212 | 84.86  | 65.71    | 310    | 21.12 |
| HHO       | 48.288506 | 49.288506 | 93.16  | 84.73    | 140    | 27.26 |
| GA        | 63.631224 | 64.631224 | 98.93  | 97.63    | 22     | 28.95 |
| HHO       | 63.572104 | 64.572104 | 99.37  | 98.60    | 13     | 41.38 |

TABLE III XGBOOST PARAMETERS OPTIMIZATION USING HHO AND GA FOR (A, C) CLASSIFICATION

|           |       |              | Class  | ificatio | on res | ults   |
|-----------|-------|--------------|--------|----------|--------|--------|
| Optimizer | N-est | L-rate       | AC (%) | k (%)    | ZOL    | To(s)  |
| HHO       | 160   | 0.2187802408 | 98.29  | 96.22    | 35     | 3.03   |
| HHO       | 140   | 0.8918738539 | 98.39  | 96.43    | 33     | 22.16  |
| GA        | 190   | 0.8197804863 | 98.44  | 96.54    | 32     | 176.07 |
| GA        | 427   | 0.5047845072 | 98.63  | 96.98    | 28     | 148.14 |

TABLE IV RF parameters optimization using HHO and GA for (A, C) Classification

|           |       | Classification results |       |     |       |  |
|-----------|-------|------------------------|-------|-----|-------|--|
| Optimizer | N-est | AC (%)                 | k (%) | ZOL | To(s) |  |
| GA        | 137   | 98.63                  | 96.97 | 28  | 4.22  |  |
| GA        | 302   | 98.78                  | 97.30 | 25  | 12.16 |  |
| HHO       | 198   | 98.83                  | 97.40 | 24  | 97.01 |  |

TABLE V K-NN PARAMETERS OPTIMIZATION USING HHO AND GA FOR (A, C) CLASSIFICATION

|           |              | Classification results |       |     |       |  |
|-----------|--------------|------------------------|-------|-----|-------|--|
| Optimizer | р            | AC (%)                 | k (%) | ZOL | To(s) |  |
| GA        | 1.9922513884 | 97.75                  | 95.03 | 46  | 0.97  |  |
| HHO       | 1.0127381746 | 98.78                  | 97.31 | 25  | 1.39  |  |
| HHO       | 2.2174709614 | 98.88                  | 97.52 | 23  | 4.69  |  |
| HHO       | 2.3406354759 | 98.93                  | 97.63 | 22  | 5.42  |  |
| HHO       | 1.1528103266 | 98.97                  | 97.74 | 21  | 7.01  |  |
| HHO       | 1.7625903302 | 99.22                  | 98.28 | 16  | 19.22 |  |
| GA        | 2.3165790702 | 99.32                  | 98.49 | 14  | 22.23 |  |
| GA        | 2.6955487125 | 99.37                  | 98.60 | 13  | 23.02 |  |
| GA        | 1.7827951227 | 99.46                  | 98.81 | 11  | 56.85 |  |

difficult to separate between alcoholics and normal, on the other hand when using the optimized filter in Table II, there is a good structure of the degrees of importance in Fig. 6.b, such that it is easy to separate between the good and bad channels, knowing that we can minimize the number of channels up to only two channels and more than reserved the best characteristics that makes it possible to identify each class. In addition, we notice that the best electrodes are positioned in the left part of the Special Issue on Smart Cities, Optimization and Modeling of Complex Systems – iJIST, ISSN : 2550-5114 Vol. 6 - No. 1 - February 2022

|            |      | Classifier performances |        |               |        |       |               |               |      |     |               |               |
|------------|------|-------------------------|--------|---------------|--------|-------|---------------|---------------|------|-----|---------------|---------------|
| Classifier | NEEG | ACC (%)                 | SE (%) | <b>PR</b> (%) | F1 (%) | K (%) | <b>MC</b> (%) | <b>JS</b> (%) | MI   | ZOL | <b>Tc</b> (s) | <b>Tp</b> (s) |
| XGB        | 61   | 99.95                   | 99.96  | 99.93         | 99.93  | 99.89 | 99.89         | 99.89         | 0.64 | 1   | 10.27         | 0.01          |
| XGB        | 10   | 99.80                   | 99.75  | 99.82         | 99.78  | 99.57 | 99.57         | 99.57         | 0.63 | 4   | 3.75          | 0.01          |
| XGB        | 5    | 98.10                   | 97.49  | 98.28         | 97.87  | 95.74 | 95.75         | 95.83         | 0.55 | 39  | 3.31          | 0.02          |
| XGB        | 4    | 95.51                   | 94.16  | 95.85         | 94.92  | 89.85 | 89.99         | 90.38         | 0.46 | 92  | 3.68          | 0.03          |
| XGB        | 3    | 90.23                   | 88.13  | 89.92         | 88.92  | 77.86 | 78.03         | 80.26         | 0.32 | 200 | 3.99          | 0.03          |
| LGBM       | 61   | 99.95                   | 99.96  | 99.93         | 99.93  | 99.89 | 99.89         | 99.89         | 0.64 | 1   | 1.51          | 0.01          |
| LGBM       | 10   | 99.76                   | 99.75  | 99.71         | 99.73  | 99.46 | 99.46         | 99.46         | 0.63 | 5   | 0.45          | 0.01          |
| LGBM       | 5    | 98.58                   | 98.24  | 98.61         | 98.42  | 96.84 | 96.85         | 96.90         | 0.57 | 29  | 0.45          | 0.01          |
| LGBM       | 4    | 96.53                   | 95.52  | 96.77         | 96.10  | 92.20 | 92.28         | 92.52         | 0.49 | 71  | 0.51          | 0.01          |
| LGBM       | 3    | 93.36                   | 91.77  | 93.37         | 92.49  | 84.99 | 85.12         | 86.14         | 0.40 | 136 | 0.86          | 0.03          |
| KNN        | 61   | 99.90                   | 99.89  | 99.89         | 99.89  | 99.78 | 99.78         | 99.78         | 0.63 | 2   | 0.22          | 1.93          |
| KNN        | 10   | 99.80                   | 99.82  | 99.75         | 99.78  | 99.57 | 99.57         | 99.57         | 0.63 | 4   | 0.04          | 0.28          |
| KNN        | 5    | 99.12                   | 98.89  | 99.16         | 99.02  | 98.04 | 98.05         | 98.06         | 0.59 | 18  | 0.04          | 0.23          |
| KNN        | 4    | 97.61                   | 97.19  | 97.49         | 97.33  | 94.67 | 94.67         | 94.82         | 0.53 | 49  | 0.02          | 0.23          |
| KNN        | 3    | 94.87                   | 93.78  | 94.77         | 94.24  | 88.49 | 88.54         | 89.18         | 0.44 | 105 | 0.02          | 0.24          |
| RF         | 61   | 99.85                   | 99.82  | 99.85         | 99.84  | 99.67 | 99.67         | 99.68         | 0.63 | 3   | 12.39         | 0.15          |
| RF         | 10   | 99.61                   | 99.46  | 99.67         | 99.57  | 99.13 | 99.13         | 99.13         | 0.62 | 8   | 5.41          | 0.14          |
| RF         | 5    | 97.51                   | 96.74  | 97.73         | 97.21  | 94.41 | 94.46         | 94.58         | 0.53 | 51  | 3.83          | 0.11          |
| RF         | 4    | 95.75                   | 94.27  | 96.31         | 95.18  | 90.37 | 90.56         | 90.85         | 0.47 | 87  | 3.92          | 0.14          |
| RF         | 3    | 92.43                   | 90.24  | 90.24         | 91.35  | 82.72 | 83.05         | 84.22         | 0.38 | 155 | 2.59          | 0.14          |
| MLP        | 61   | 99.71                   | 99.64  | 99.71         | 99.67  | 99.35 | 99.35         | 99.35         | 0.62 | 6   | 14.67         | 0.05          |
| MLP        | 10   | 99.61                   | 99.57  | 99.57         | 99.57  | 99.13 | 99.13         | 99.14         | 0.62 | 8   | 28.08         | 0.03          |
| MLP        | 5    | 96.29                   | 94.85  | 96.96         | 95.79  | 91.59 | 91.79         | 91.95         | 0.49 | 76  | 70.28         | 0.05          |
| DNN        | 61   | 99.56                   | 99.53  | 99.50         | 99.51  | 99.02 | 99.02         | 99.03         | 0.61 | 9   | 20.15         | 0.03          |
| DNN        | 10   | 99.76                   | 99.71  | 99.75         | 99.73  | 99.46 | 99.46         | 99.46         | 0.63 | 5   | 28.76         | 0.07          |
| DNN        | 5    | 94.63                   | 92.80  | 95.28         | 93.88  | 87.77 | 88.05         | 88.54         | 0.44 | 110 | 56.71         | 0.02          |
| DNN        | 4    | 89.26                   | 86.33  | 89.52         | 87.60  | 75.28 | 75.78         | 75.23         | 0.30 | 220 | 58.75         | 0.02          |

 TABLE VI

 CLASSIFICATION RESULTS FOR ALCOHOLISM DATASET BINARY-CLASS USING MACHINE LEARNING ALGORITHMS

central lobe and in the right part of the frontal and parietal lobes, which implies that these areas are more sensitive to alcoholism and consequently that all the nervous systems of the organs related to these parts of the brains will be negatively influenced.

## C. Classification Results

Table III shows optimization of the EEG signals classification when applying GA algorithm on the XGB parameters, such that the employing of the bandpass filter can increase the accuracy value to over 97% and the GA algorithm looks for efficient and stable combinations of XGB parameters and the same procedure for RF in Table IV and for k-NN in Table V, so this table shows an accuracy value increase over of 98.7% and a decrease of ZOL value from 118 to 52 errors. Thus, when the number of estimators increases, also the classification time and accuracy values increase, but the most important in an EEG signal prediction system are the classification speed, the prediction speeds, the accuracy value, and the number of electrodes used for EEG acquisition. Table VI shows a classification of alcoholism EEG dataset when using a bandpass filter with order 3 and bandwidth between 63.572 and the number of electrodes decreases to 10, the accuracy values remain more than 99.5%, which shows a good state of prediction system stability also when the number of the electrodes decreases. The prediction speed of this system remains very fast between 25501 and 76923 samples per second, so the classification speed is also between 408 and 4625 samples per second, this speed shows the goodness of the system presented in this paper. Also, this table shows that the accuracy value remains high at the order of 95% during the third condition using only four of the most contributed electrodes (P8, FC4, FZ, and PZ), these electrodes characterize a brain activity variation at the level of both frontal and parietal lobes at the extremity of the left and right temporal lobes. This table shows classification results using other classifiers like RF, KNN, MLP, and DNN. Knowing that the maximum accuracy value is 99.90% found using k-NN but with a bad prediction speed of 106 samples per second, against XGB allows predicted this EEG data with a speed of 33820 samples per second, i.e. a speed 300 times higher than the speed when using the k-NN

64.572Hz, such that all the accuracy values are more than

99.5%, which shows the filter used efficiency, as well as when

Special Issue on Smart Cities, Optimization and Modeling of Complex Systems – iJIST, ISSN : 2550-5114 Vol. 6 - No. 1 - February 2022

| TABLE VII  |
|--|
| OVERVIEW OVER OTHER WORKS PERFORMING BINARY CLASSIFICATION |
| RESULTS FOR ALCOHOLISM DETECTION DATASET                   |

| Work                     | AC      | Methods                                 |  |  |  |  |
|--------------------------|---------|---|--|--|--|--|
| Acharya et al. [7]       | 91.7 %  | SVM                                     |  |  |  |  |
| Anuragi and Sisodia [10] | 97.06 % | Tunable Q-wavelet transform with LS-SVM |  |  |  |  |
| Anuragi and Sisodia [11] | 99.17%  | LS-SVM                                  |  |  |  |  |
| Anuragi and Sisodia [12] | 98.75%  | EWD with LS-SVM                         |  |  |  |  |
| Ehlers et al. [23]       | 88%     | CD with Discriminant analysis           |  |  |  |  |
| Faust et al. [26]        | 93.33%  | PDF extraction with SVM                 |  |  |  |  |
| Faust et al. [27]        | 95.8%   | WPD, energy measures with               |  |  |  |  |
|                          |         | kNN                                     |  |  |  |  |
| Mehla et al. [45]        | 99.98%  | FDM-LDA                                 |  |  |  |  |
| Mumtaz et al. [47]       | 96%     | LMT                                     |  |  |  |  |
| Mumtaz et al. [48]       | 98%     | SVM                                     |  |  |  |  |
| Patidar et al. [52]      | 97.02%  | Tunable Q-wavelet transform with LS-SVM |  |  |  |  |
| Rachman et al. [54]      | 85%     | Daubechies-PNN                          |  |  |  |  |
| Rodrigues et al. [56]    | 99.60%  | Symlets with Nave-Bayes                 |  |  |  |  |
|                          | 99.87%  | Biorthogonal with Nave-<br>Bayes        |  |  |  |  |
| Taran and Bajaj [63]     | 97.92%  | LS-SVM                                  |  |  |  |  |
| Upadhyay et al. [64]     | 100%    | Morlet-RF                               |  |  |  |  |
| This work                | 99.95%  | Gamma-XGB                               |  |  |  |  |
|                          | 99.90%  | Gamma-KNN                               |  |  |  |  |
|                          | 99.76%  | Gamma- <b>R</b> F                       |  |  |  |  |

algorithm. Table VII presents the classification results for other methods when classifying EEG signals from the alcoholics and normal subjects, knowing that the best accuracy value found is 100% found by Upadhyay et al. [64] using a Morlet-RF method, but this work does not present the calculation procedure to properly show its results. The second-best method finds an accuracy value of 99.98% found by Mehla et al. [45] using FDM-LDA, this method shows a good selection of orthogonal features between them to increase the independence index and then minimize the correlation between all electrodes. The method used by Rodrigues et al. [56] also finds a good precise value of 99.87% using the algorithm Nave-Bayes and the decomposition of data using Biorthogonal, this method is based on the same principle of the FDMLDA method for selecting orthogonal features between them, but these methods are very slow during a real-time for classification and prediction stages, because of more load of the decomposition algorithms used compared to our method which is based on a simple bandpass filter and the XGB algorithm for data classification with accuracy values more than 99% using more than 4 electrodes to acquire EEG signals. The work of Anuragi and Sisodia [12] finds a good accuracy value of 98.73% when applying EWT to extract EEG waves, such that the conversion of waves using HHT can extract a set of characteristics and with the use of the LS-SVM algorithm for classification shows that this method is preferable, but the use of features extracted in real-time can minimize the prediction speed of the global system in time real compared to our method because of the high density of calculation at each moment. Also, the work of Mumtaz et al. [48] uses the statistical derivations of raw EEG and the extraction of the characteristics using SL and ROC function to minimize the number of features that have been extracted up to the minimum possible, which gives as results an accuracy value of the order of 98% which is a minimal can be compared to that found by Anuragi and Sisodia [12] probably because of the use of basic classify as SVM, NB, logistic regression (LR). In the work of Taran and Bajaj [63] we can also notice that the precision value was found is at the level of 97.92% using low frequencies extracted using the IMF function, such that it extracts from these frequencies some characteristics like entropy, mean, variance, etc. to increase the degree of separation between the two classes, this method shows the results of the optimization of Table II, as illustrated in Fig. 4 that the frequency of signals increases in the case of alcoholics compared to normal people, which shows that it is possible to use low frequencies for the classification shown by this work of Taran and Bajaj [63] or by using high frequencies of the gamma wave in our work. Also, the work of Patidar et al. [52] presents the results of the application of TQWT for the extraction of EEG signal characteristics and the use of ICA filter in combination with the LS-SVM classification to minimize the correlation values between features and improved data classification, such that this method shows that the accuracy value is at the level of 97.02% which is also a good prediction but it remains not ideal by compared to other methods

#### V. CONCLUSION

In conclusion, this work shows the great performance of the machine learning algorithms for the Alcoholism EEG classification, such that the use of the bandpass filter can eliminate all existing links between all electrodes and extract only the best features that are found in low frequency, the employing of correlation matrix makes it easy to see the effect of the filter, the machine learning algorithms also shows its classification quality with high pus accuracy values of 99% and prediction speeds passes 25500 samples per second, knowing that the application of the third condition on subjects can easily deduce alcoholism compared to normal with an accuracy value of 100%, this value summarizes the alcoholism detection quality using simple and fast method compared to the literature results. It is hoped that this work can help other biologists and medical researchers. Our future work is focused on real applications in diagnostic and control based on EEG signals.

#### REFERENCES

- S. Abenna, M. Nahid, and A. K. Belbachir. Brain-computer interface: Rhythm alpha analysis for eyes signals. In The Fourth International Conference On Intelligent Computing in Data Sciences, IEEE., 2020. doi: 10.1109/ICDS50568.2020.9268719.
- [2] S. Abenna, M. Nahid, and A. Bajit. BCI: Classifiers Optimization for EEG Signals Acquiring in Real-Time. In 2020 6th IEEE Congress on Information Science and Technology (CiSt), 2021. doi: 10.1109/CiSt49399.2021.9357209.
- [3] S. Abenna, M. Nahid, and A. Bajit. Brain-computer interface: A novel EEG classification for baseline eye states using LGBM algorithm, volume 211, chapter 18, pages 189–198. Springer Science and Business Media LLC, 2021. ISBN 978-3-030-73882-2. doi: 10.1007/978-3-030-73882-2 18.
- [4] S. Abenna, M. Nahid, and A. Bajit. Motor imagery based brain-computer interface: improving the eeg classification using delta rhythm and lightgbm algorithm. Biomedical Signal Processing and Control, 71(103102), January 2022. doi: 10.1016/j.bspc. 2021.103102.
- [5] U. R. Acharya. Nonlinear dynamics measures for automated eeg-based sleep stage detection. Eur. Neurol., 2015.

- [6] U. R. Acharya, E. C. Chua, K. C. Chua, L. C. Min, and T. Tamura. Analysis and automatic identification of sleep stages using higher order spectra. Int. J. Neural Syst., 2010.
- [7] U. R. Acharya, S. Sree, S. Chattopadhyay, and J. SURI. AUTOMATED DIAGNOSIS OF NORMAL AND ALCOHOLIC EEG SIGNALS. *Journal of Neural Systems*, 22(3):1–11, 2012. doi: 10.1142/S0129065712500116.
- [8] N. Ahmadi, Y. Pei, and M. Pechenizkiy. Detection of alcoholism based on eeg signals and functional brain network features extraction. *In 2017 IEEE 30th International Symposium on Computer-Based Medical Systems (CBMS), IEEE*, 2017. doi: 10.1109/CBMS.2017.46.
- [9] E. Alickovic and A. Subasi. Ensemble SVM Method for Automatic Sleep Stage Classification. pages 1–8, 2018.
- [10] A. Anuragi and D. S. S. Sisodia. Alcoholism detection using support vector machines and centered correntropy features of brain eeg signals. *In* 2017 International Conference on Inventive Computing and Informatics (ICICI), IEEE, 2017. doi: 10.1109/ICICI.2017.8365294.
- [11] A. Anuragi and D. S. S. Sisodia. Alcohol use disorder detection using eeg signal features and flexible analytical wavelet transform. *Biomedical Signal Processing and Control*, 2019. doi: 10.1016/j.bspc.2018.10.017.
- [12] A. Anuragi and S. S. S. Sisodia. Empirical wavelet transform based automated alcoholism detecting using eeg signal features. *Biomedical Signal Processing and Control*, 2020. doi: 10.1016/j.bspc.2019.101777.
- [13] A. Asghar, S. Mirjalili, H. Faris, and I. Aljarah. Harris hawks optimization : Algorithm and applications. Future Generation Computer Systems, 97: 849–872, 2019. ISSN 0167-739X. doi: 10.1016/j. future.2019.02.028.
- [14] M. Buhler and K. Mann. Alcohol and the human "brain: a systematic review of different neuroimaging methods. Alcohol Clin. Exp. Res., 2011. doi: 10.1111/j.1530-0277.2011.01540.x.
- [15] A. Brandt. A signal processing framework for operational modal analysis in time and frequency domain. *Mech. Syst. Signal Process*, 2019. doi: 10.1016/j.ymssp.2018.06.009.
- [16] L. Breiman. Random forests. *Machine Learning*, 2001. doi: 10.1023/A:1010933404324.
- [17] J. Camprodon and T. Stern. Selecting neuroimaging techniques: A review for the clinician. The Primary Care Companion for CNS Disorders, 2012. doi: 10.4088/PCC.12f01490.
- [18] T. Chen and C. Guestrin. Xgboost: A scalable tree boosting system. In Proceedings of the 22nd Acm Sigkdd International Conference on Knowledge Discovery and Data Mining, page 785–794, 2016. doi: 10.1145/2939672.2939785.
- [19] B. Chowdhury and G. Garai. A review on multiple sequence alignment from the perspective of genetic algorithm. *Genomics*, 109: 419–31, 2017. doi: 10.1016/j.ygeno.2017.06.007.
- [20] S. Churchill and L. Farrell. Alcohol and depression: evidence from the 2014 health survey for england. *Drug Alcohol Depend*, 180:86–92, 2017. doi: 10.1016/j.drugalcdep.2017.08.006.
- [21] D. Coit. Genetic algorithms and engineering design. Eng. Econ, 43:379– 81, 1998. doi: 10.1080/00137919808903206.
- [22] L. G. Doroshenkov, V. A. Konyshev, and S. V. Selishchev. Classification of human sleep stages based on eeg processing using hidden markov models. Biomed. Eng., 2007.
- [23] C. Ehlers, J. Havstad, D. Prichard, and J. Theiler. Low doses of ethanol reduce evidence for nonlinear structure in brain activity. J. Neurosci., 18:7474–7486, 1998.
- [24] M. Enoch and D. Goldman. Problem drinking and alcoholism: diagnosis and treatment. Am. Fam. Phys., 65:441–449, 2002.
- [25] S. Enshaeifar, S. Kouchaki, C. C. Took, and S. Sanei. Quaternion singular spectrum analysis of electroencephalogram with application in sleep analysis. IEEE Trans. Neural Syst. Rehabil. Eng., 2016.
- [26] O. Faust, R. Acharya, A. Allen, and C. Lin. Analysis of eeg signals during epileptic and alcoholic states using ar modeling techniques. IRBM, 29:44–52, 2008. doi: 10.1016/j.rbmret.2007.11.003.
- [27] O. Faust, W. Yu, and N. Kadri. Computer-based identification of normal and alcoholic eeg signals using wavelet packets and energy measures. J. Mech. Med. Biol., 13, 2013. doi: 10.1142/S0219519413500334.
- [28] P. Geurts, D. Ernst, and L. Wehenkel. Extremely randomized trees. Machine Learning, 2006.
- [29] M. Gong, Y. Bai, J. Qin, J. Wang, P. Yang, and S. Wang. Gradient boosting machine for predicting return temperature of district heating system: A case study for residential buildings in tianjin. *Journal of Building Engineering*, 27, 2020. doi: 10.1016/j.jobe.2019.100950.
- [30] V. Gupta and M. Mittal. KNN and PCA classifier with Autoregressive modelling during different ECG signal interpretation. *Procedia Computer*

*Science*, 125:18–24, 2018. ISSN 18770509. doi: 10.1016/j.procs.2017.12.005.

- [31] C. M. Held. Extracting fuzzy rules from polysomnographic recordings for infant sleep classification. IEEE Trans. Biomed. Eng., 2006.
- [32] J. Holland. Genetic algorithms. *Sci Am*, 267: 66–72, 1992.
- [33] H. Houssein, M. E. Hosney, D. Oliva, W. M. Mohamed, and M. Hassaballah. A Novel Hybrid Harris Hawks Optimization and Support Vector Machines for Drug Design and Discovery. *Computers and Chemical Engineering*, page 106656, 2019. ISSN 0098-1354. doi: 10.1016/j.compchemeng.2019.106656.
- [34] E. Ianez, J. Azorn, A. Ubeda, J. Ferrandez, and E. Fernandez. Mental tasks-based brainrobot interface. *Robotics and Autonomous Systems*, 58(12):1238–1245, 2010. doi: 10.1016/j.robot.2010.08.007.
- [35] H. Jiajun, Y. Chuanjin, L. Yongle, and X. Huoyue. Ultra-short term wind prediction with wavelet transform, deep belief network and ensemble learning. *Energy Conversion and Management*, 205, 2020. doi: 10.1016/ j.enconman.2019.112418.
- [36] D. Jiang, Y.-n. Lu, Y. Ma, and Y. Wang. Robust sleep stage classification with singlechannel EEG signals using multimodal decomposition and HMM-based refinement. 121:188–203, 2019. doi: 10.1016/j.eswa.2018. 12.023.
- [37] D. Jiang, Y. MA, and Y. Wang. Sleep stage classification using covariance features of multi-channel physiological signals on riemannian manifolds. *Computer Methods and Programs in Biomedicine*, 178:19–30, 2019. doi: 10.1016/j.cmpb.2019.06.008.
- [38] H. G. Jo, J. Y. Park, C. K. Lee, S. K. An, and S. K. Yoo. Genetic fuzzy classifier for sleep stage identification. Comput. Biol. Med., 2010.
- [39] K. Kalaiselvi, K. Sangeetha, and S. Mogana. Efficient Disease Classifier Using Data Mining Techniques: Refinement of Random Forest Termination Criteria. Journal of Computer Engineering, 14(5): 104–111, Sep 2013. ISSN 2278-0661.
- [40] I. Kononenko. Machine learning for medical diagnosis: History, state of the art and perspective. Artif. Intell. Med., 2001.
- [41] S. F. Liang, Y. H. Kuo, Y. H. Hu, Y. H. Pan, and Y. H. Wang. Automatic stage scoring of single-channel sleep eeg by using multi-scale entropy and autoregressive models. IEEE Trans. Instrum. Meas., 2012.
- [42] A. Liaw and M. Wiener. Classification and regression by random forest. R News, 2002.
- [43] S. Lim and S. Chi. Advanced Engineering Informatics Xgboost application on bridge management systems for proactive damage estimation. Advanced Engineering Informatics, 41, May 2019. ISSN 1474-0346. doi: 10.1016/j.aei.2019.100922.
- [44] J. Lin, H. Chen, S. Li, Y. Liu, X. Li, and B. Yu. Accurate prediction of potential druggable proteins based on genetic algorithm and bagging-svm ensemble classifier. *Artificial Intelligence In Medicine*, 98:35–47, 2019. doi: 10.1016/j.artmed.2019.07.005.
- [45] V. K. Mehla, A. Singhal, and P. Singh. A novel approach for automated alcoholism detection using fourier decomposition method. *Neuroscience Methods*, 2020. doi: 10.1016/j.jneumeth.2020.108945.
- [46] G. Mohammadi, P. Shoushtari, B. Ardekani, and M. Shamsollahi. Person identification by using ar model for eeg signals. World Academy of Science, Engineering and Technology, pages 281–285, 2006. ISSN 1307-6884.
- [47] W. Mumtaz, P. Vuong, L. Xia, A. Malik, and R. Rashid. Automatic diagnosis of alcohol use disorder using eeg features. *Knowledge-Based Systems*, pages 48–59, 2016. doi: 10.1016/j. knosys.2016.04.026.
- [48] W. Mumtaz, M. N. M. Saad, N. Kamel, S. S. A. Ali, and A. S. Malik. An eeg-based functional connectivity measure for automatic detection of alcohol use disorder. *Artificial Intelligence in Medicine*, pages 79–89, 2017. doi: 10.1016/j.artmed.2017.11.002.
- [49] W. Mumtaz, P. Vuong, A. S. Malik, and R. Rashid. A review on eegbased methods for screening and diagnosing alcohol use disorder. *Cogn. Neurodyn.*, 2018.
- [50] M. Nicolelis. Actions from thoughts. *Nature*, 409:403–407, 2001. doi: 10.1038/35053191.
- [51] W. Organization. Alcohol key facts. Technical Report, 2018.
- [52] S. Patidar, R. Pachori, A. Upadhyay, and U. Acharya. An integrated alcoholic index using tunable-q wavelet transform based features extracted from eeg signals for diagnosis of alcoholism. *Applied Soft Computing*, 2017. doi: 10.1016/j.asoc.2016.11.002.
- [53] M. Poulos, M. Rangoussi, and E. Kafetzopoulos. Person identification via the eeg using computational geometry algorithms. In 6<sup>th</sup> IEEE International conference on Electronics, Circuits and Systems, pages 1005–1008, 1999. ISBN 978-960-7620-06-4.

- [54] N. Rachman, H. Tjandrasa, and C. Fatichah. Alcoholism classification based on eeg data using independent component analysis (ica), wavelet de-noising and probabilistic neural network (pnn). *In 2016 International Seminar on Intelligent Technology and Its Applications (ISITIA), IEEE*, 2016. doi: 10.1109/ISITIA. 2016.7828626.
- [55] C. E. Rasmussen and C. K. I. Williams. Gaussian processes for machine learning. Cambridge, MA, USA: MIT, 2006.
- [56] J. d. C. Rodrigues, P. P. Filho, E. Peixoto, A. K. N, and V. H. C. de Albuquerque. Classification of EEG signals to detect alcoholism using machine learning techniques. *Pattern Recognition Letters*, 125:140–149, 2019. ISSN 01678655. doi: 10.1016/j.patrec.2019.04.019.
- [57] R. Salazar-Varas and R. A. Vazquez. Evaluating the effect of the cutoff frequencies during the pre-processing stage of motor imagery eeg signals classification. *Biomedical Signal Processing and Control*, 54, 2019. doi: 10.1016/j.bspc.2019.101592.
- [58] M. Schuckit. Drug and alcohol abuse: a clinical guide to diagnosis and treatment. *Springer Sci. Bus. Media*, 2006.
- [59] M. Schuckit. Recognition and management of withdrawal delirium (delirium tremens). N. Engl. J. Med., 2014. doi: 10.1056/ NEJMra1407298.
- [60] M. R. Silvestre and L. L. Ling. Pruning methods to MLP neural networks considering proportional apparent error rate for classification problems with unbalanced data. *Measurement: Journal of the International Measurement Confederation*, 56:88–94, 2014. ISSN 02632241. doi: 10.1016/j.measurement.2014. 06.018.
- [61] M.-Y. Su. Using clustering to improve the knn-based classifiers for online anomaly network traffic identificatio. *Journal of Network and Computer Applications*, 34:722–730, 2011. doi: 10.1016/j.jnca.2010.10.009.
- [62] A. Subasi, S. Jukic, and J. Kevric. Comparison of emd, dwt and wpd for the localization of epileptogenic foci using random forest classifie. *Measurement*, 146:846–855, 2019. doi: 10.1016/j.measurement.2019. 07.026.
- [63] S. Taran and V. Bajaj. Rhythm-based identification of alcohol eeg signals. Measurement Technology, 2018. doi: 10.1049/iet-smt.2017.0232.
- [64] R. Upadhyay, P. Padhy, and P. Kankar. Alcoholism diagnosis from eeg signals using continuous wavelet transform. *In 2014 Annual IEEE India Conference (INDICON), IEEE*, 2014. doi: 10.1109/INDICON.2014. 7030476.
- [65] W. H. O. (WHO). In 2014. Global Status Report on Alcohol and Health (Online).
- [66] T. Willemen. An evaluation of cardiorespiratory and movement features with respect to sleep-stage classification. IEEE J. Biomed. Health Inform., 2014.
- [67] H. T. Wu, R. Talmon, and Y. L. Lo. Assess sleep stage by modern signal processing techniques. IEEE Trans. Biomed. Eng., 2015.
- [68] G. Zhu, Y. Li, and P. P. Wen. Analysis and classification of sleep stages based on difference visibility graphs from a single-channel eeg signal. IEEE J. Biomed. Health Informat., 2014.