BIOCYBERNETICS AND BIOMEDICAL ENGINEERING XXX (2019) XXX–XXX

Available online at www.sciencedirect.com
ScienceDirect
journal homepage: www.elsevier.com/locate/bbe

Original Research Article

A machine learning approach to epileptic seizure prediction using Electroencephalogram (EEG) Signal

Q2 Marzieh Savadkoohi [a], Timothy Oladduni [b,*]

[a] School of Eng and Applied Sciences, University of District of Columbia, Washington DC, USA
[b] Department of Computer Science, University of District of Columbia, Washington DC, USA

ARTICLE INFO

Article history:
Received 8 November 2019
Received in revised form
17 June 2020
Accepted 7 July 2020
Available online xxx

Keywords:
Electroencephalogram (EEG)
Epileptic seizure
K-Nearest neighbors (KNN)
Support vector machines (SVM)
Time-Domain
Frequency-Domain
Wavelet transform

ABSTRACT

This study investigates the properties of the brain electrical activity from different recording regions and physiological states for seizure detection. Neurophysiologists will find the work useful in the timely and accurate detection of epileptic seizures of their patients. We explored the best way to detect meaningful patterns from an epileptic Electroencephalogram (EEG). Signals used in this work are 23.6 s segments of 100 single channel surface EEG recordings collected with the sampling rate of 173.61 Hz. The recorded signals are from five healthy volunteers with eyes closed and eyes open, and intracranial EEG recordings from five epilepsy patients during the seizure-free interval as well as epileptic seizures. Feature engineering was done using; i) feature extraction of each EEG wave in time, frequency and time-frequency domains via Butterworth filter, Fourier Transform and Wavelet Transform respectively and, ii) feature selection with T-test, and Sequential Forward Floating Selection (SFFS). SVM and KNN learning algorithms were applied to classify preprocessed EEG signal. Performance comparison was based on Accuracy, Sensitivity and Specificity. Our experiments showed that SVM has a slight edge over KNN.

1. Introduction

Human brain serves as the most important part of the central nervous system. It composes of billions of cells that are mostly neurons. Each neuron is made up of axons, dendrites, and cellular bodies. They react to the stimulus and transmit information in long paths to other neurons and organs such as muscles and gland cells [1]. Dendrites are connected to axons or dendrites of other neurons and receive or transmit impulses to them. The main type of communication between billions of neurons in the human brain is dendritic communication. Therefore, the brain is a very complicated interconnected network of neurons. These neurons are stimulated at any moment and produce electric fields [2]

The time-varying electrical currents produced by neurons at the cell membrane surface are from two types of neuronal activities: a) rapid depolarization of neuronal membranes which depends on the voltage of sodium and potassium ions and results in the Action Potential (AP) [3]. b) Slower changes of

* Corresponding author.
E-mail addresses: marzieh.savadkoohi@udc.edu (M. Savadkoohi), timothy.oladunni@udc.edu (T. Oladduni).
https://doi.org/10.1016/j.bbe.2020.07.004

membrane potential due to the synaptic activity and the function of several neurotransmitter systems [4]. AP is a rapid change, 1–2 milliseconds, in the membrane potential, which changes the intracellular potential from negative to positive and returns rapidly to the intracellular resting potential. It has smaller field potential distribution (less penetration in the extracellular environment) and is less sustained (about 1 millisecond compared to post-synaptic potentials extending from 15 to more than 200 milliseconds) [3,5]. Post-synaptic potential spreads out to the skull surface and can be measured.

To measure the brain activity, we can use a non-invasive method of putting electrodes on the scalp through a device called Electroencephalogram (EEG).EEG signal are mainly produced by the measurable potential of post-synaptic pyramidal cells, which are parallel to each other and perpendicular to the skull's surface. It creates an extracellular cortical dipolar layer [5,6]. Therefore, the electrodes on the skull represent the time and place of the post-synaptic potential of the cortical neurons. It also includes slow and simultaneous potential changes in the large cortical regions (Fig. 1) [7,8].

EEG signal are useful in identifying many clinical problems such as schizophrenia, Alzheimer, insomnia, sleep disorders, seizure disorders, brain tumors and infections of the central nervous system. Besides being non-invasive and having exquisite temporal resolution, this technique provides low cost and needs no extreme safety restrictions [5]. Using EEG signal, it is evident that epileptic seizures usually start spontaneously. They result from sudden electrical discharge of part of the brain cells and therefore cause temporary agitation of the brain. Sometimes seizures might go disregarded or may be confused with other brain disorders such as a meningitis or stroke which can also cause the same symptoms.

Study shows that approximately, 10% of people experience at least one seizure in their lifetime [10]. Precise analysis of epileptic in the electroencephalograph (EEG) signal can reveal valuable facts about this prevalent brain disorder [11]. Since EEG signal is very complex, it therefore requires the analysis of several factors. Manual visual inspection of EEG signal has been found useful in identifying patterns. However, this approach requires high level of technical and analytical skills with several signal-processing techniques [12]. Therefore, in the recent times, automating epileptic seizure detection has gained attention among researchers.

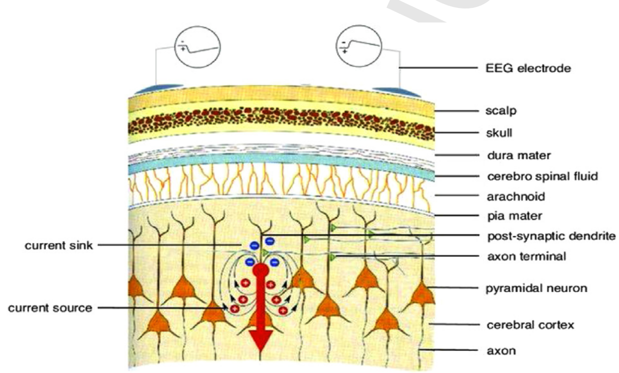

Fig. 1 – Diagram of neurons firing during EEG [9].

Digital recording, saving and analyzes of EEG signal have become possible by the advancement of technology during the past few decades. This digitized data can be given to an automated seizure detection system in order to detect seizures found in the EEG data [13]. EEG Signal analysis includes three main steps: preprocessing, feature extraction (plus feature selection), and classification. Preprocessing stage is mainly about signal acquisition, noise removal, averaging, thresholding, signal enhancement, etc. After this stage, the markers are processed using suitable feature extraction methodologies to find the most informative features. With the help of feature selection module, which is at an optional stage, we can reduce the vector size. Therefore, the foremost related features needed for discrimination are found. Classification module is the final step. It tests the feature vector and finds the best way to classify the attributes based on the hidden pattern that govern them and therefore their algorithmic model [14].

In this study, we explored machine learning approach in the diagnoses and detection of epileptic seizure. Doing this, we used K-nearest neighbors (KNN) and Support Vector Machine (SVM) classification systems. We analyzed the EEG signal in the time and frequency domains and extracted the features of each EEG wave. After comparing the efficiency of each feature individually, we combined the features together and chose the best and most significant features using two different techniques: T-test and Sequential Forward Floating Selection (SFFS). The best selected features were then fed into SVM and KNN classifiers and the results of our models were evaluated using k- fold cross-validation methodology.

The remaining parts of the paper is organized as follows; Section 2 is the review of related research work on the recording EEG data. We give a brief description of the dataset in section 3. The proposed methodology is in section 4. Section 5 is the experimental result showing the effectiveness of our method. Finally, we conclude this paper in Section 6 with the implication of the study in section 7.

2. Literature review

Typically, an EEG test is recorded in a rest position (open/closed eyes) or during a specific task, such as listening, reading, watching, or calculating something [15]. Then, the computerized analysis of the EEG signal investigates several factors including: signal frequency distribution, EEG signal amplitudes, spatial coordinates of specific phenomena occurrence, morphology of the waveform, waveform for similar regions of the brain, symmetry between the brain hemispheres (voltage symmetry, frequency symmetry), the occurrence model of the waveform (random, sequential, continuous) and reactivity (change in the individual's state and subsequent changes in an EEG parameter). Afterwards, the most important step is comparing the calculated components for the subject with the normative database. First, artifacts should be correctly removed, and the used normal data should be consistent with the subject's age, gender and conditions. Then the results achieved for the subject under study is compared with the normal Z-scores [16].

Several studies have been done on the classification of EEG signal. Different researchers have applied variety of methods

Please cite this article in press as: Savadkoohi M, Oladduni T. A machine learning approach to epileptic seizure prediction using Electroencephalogram (EEG) Signal. Biocybern Biomed Eng (2020), https://doi.org/10.1016/j.bbe.2020.07.004

to analyze EEG records for research and clinical applications and reported their results. These studies used different techniques and achieved reasonable classification accuracies for the same dataset that we have used in our work. This dataset was produced by Andrzezak et al. [17] and downloaded from epileptologie-bonn webpage [18].

Among different studies, Oweis et al. used Hilbert-Huang Transform to extract the EEG signal frequency features and 94% accuracy was determined through the Empirical Mode Decomposition (EMD) classification method. Their results also showed 92% sensitivity and 96% specificity. EMD is a well-known method that performs time-frequency analysis on data with non-linear or non-stationary nature such as EEG signal. [19]. Sharma et al. did discrimination between different classes of the same EEG data through Least-squares support vector machine (LS-SVM). They achieved a maximum accuracy of 100% for the classification of healthy versus epileptic patients [20].

Acharya and his colleagues applied a deep convolutional neural network (CNN) composed of 13 layers to analyze the EEG data. They achieved total accuracy of 88.67%, specificity of 90% and sensitivity of 95% to characterize people as normal, preictal and seizure groups [21]. Another group who worked on the same dataset were Guha and his co-workers. They proposed a Deep Neural Network (DNN) system, composed of five hidden layers, to detect epileptic seizure and compared that with traditional classifiers such as multilayer perceptron (MLP) and k-nearest neighbor (K–NN). Their results showed 80% accuracy, 80% sensitivity and 64% precision in predicting epileptic seizure with DNN which was higher than KNN and MLP performance (76% and 78% accuracy respectively) [22].

For automated diagnosis of epilepsy, Patidar et al. used tunable-Q wavelet transform (TQWT) method and decomposed EEG signal to several sub-bands. Then they extracted Kraskov entropy as the main and only characteristic to obtain the value of non-linearity in the EEG signal. Next, they applied Least Square Support Vector Machine (LS-SVM) to classify seizure versus seizure-free signal and got 97.75% of accuracy, 97% sensitivity and 99% specificity [23].

To distinguish between ictal and non-ictal EEG signal, Zahra et al. employed Multivariate Empirical Mode Decomposition (MEMD) algorithm [24]. This technique is an extension of EMD method which was used by Oweis et al. and explained earlier [19]. After selecting necessary features in the time-frequency domain they used neural networks and achieved overall 87.2% accuracy [24].

In another work, Bhattacharyya et al. extracted entropy from different EEG signal frequency sub-bands. In order to do so, they firstly decomposed EEG signal with Q-tunable wavelet transform method. Then used K-Nearest Neighbors (K–NN) to extract entropies from various sub-bands of interest. Finally, they fed the extracted features to the SVM machine and got accuracies in the range of 98%–100% for classification of different categories, e.g. normal versus epileptic subjects, seizure versus seizure free signal, etc. [25]

Richhariya et al. employed universum support vector machine (USVM) to detect epileptic signals and differentiate them from normal brain signals in healthy people. The method they used takes the advantage of removing outliers and their effect on the generation of universum data. Also, to gain less computational cost compared to traditional SVMs, they applied universum twin support vector machine (UTSVM) strategy. Among different classes their results showed highest accuracy of 99% for classification of healthy and epileptic EEG signals and higher generalization performance compared to traditional SVM and its different derivatives such as USVM, Twin SVM (TWSVM) and UTSVM [26].

Kaya used Minimum Redundancy Maximum Relevance (mRMR) feature selection method to reduce the size of extracted features and keep the most relevant attributes. Then fed the selected features to the fine and weighted K-Nearest Neighbors (k-NN) models and achieved 98.78 %, 98.56 %, respectively [27].

The dataset [18] for this work was preprocessed and cleaned of artifacts. It is remarkably popular among researchers of different disciplines such as: neuroscience, biomedical sciences, medicine, computer science and other fields. A quick look at literature, shows lots of works on the same data, each with advantages and disadvantages. In this wok, we aim to distinguish between healthy and epileptic patients, with the highest possible accuracy, using two popular machine learning techniques, namely, SVM and KNN in three time, frequency and time-frequency domains simultaneously. Extracting same statistical features from five brain waves in all three different domains, keeps data exhaustive while providing an equal analytical condition to compare the performance of the classifiers.

3. Dataset description

The dataset we used to study the epileptic seizures was produced by Andrzezak et al. at university of Bonn [17]. it is publicly available at [18]. This data contains EEG signal of five healthy participants as well as five patients who were diagnosed with epilepsy. Two resting situations of eyes open and eyes closed were used to record the brain EEG signal for healthy subjects. The standard 10–20 scheme was used to place the electrodes on the subjects' scalp and the signal recorded continuously. The dataset consists of 5 folders (A–E) and each folder is made of 100 single channel EEG segments which was recorded at the sampling rate of 173.61 Hz and band-pass filter setting of 0.53–40 Hz.

Although the original recorded dataset was a multi-channel continuous data, only some segments of this dataset were taken. Artifacts from breathing, eye movement, muscle activity, etc. Folders A, B, C, D and E were created. Folders A and B contained information of the healthy participants in eyes-open and eyes-closed situation respectively, while C, D and E represent the EEG signal of the epileptic patients. More specifically, folder C and D are the EEG signal during the seizure free intervals of the patients who had complete seizure control after the epileptogenic zone resection. The recorded signal in set C were from epileptogenic zone and signal in set D were recorded from the same region in the opposite brain hemisphere. Set E is the only one which contains signal during the seizure activity [17,18]. Fig. 3 shows one sample EEG signal from each subset A, B, C, D and E respectively.

Please cite this article in press as: Savadkoohi M, Oladduni T. A machine learning approach to epileptic seizure prediction using Electroencephalogram (EEG) Signal. Biocybern Biomed Eng (2020), https://doi.org/10.1016/j.bbe.2020.07.004

4. Methodology

The amplitude of the brain waves recorded at the skull surface is very weak (about 0–100 microvolts with a frequency of about 0.5–100 Hz). The brain waves can be classified according to their frequencies into the following five categories: Delta (0.1–4 Hz), Theta (4–8 Hz), Alpha (8–12 Hz), Beta (12–30 Hz) and Gamma (30–100 Hz) [5,28]. In pattern recognition, we usually reduce the vector size to a set of selected features, rather than the main signal which may take a huge space or be dependent on external conditions. As a result, we kept the most important and relevant features and removed the redundant attributes [29,30]. The architectural diagram of the proposed model is shown in Fig. 2.

Features were extracted based on the following statistical characteristics; mean, variance, skewness and kurtosis, in the time, frequency and time-frequency domains [31]. The mathematical formulations of mean, variance, skewness and kurtosis are shown in Eq. (1–4):

$$Mean : E(x) = \frac{1}{n}\sum_{i=1}^{n} x_i \tag{1}$$

$$Variance : Var(x) = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - E(x))^2 \tag{2}$$

$$Skewness : S(x) = E\left[\left(\frac{x - E(x)}{\sqrt{Var(x)}}\right)\right]^3 \tag{3}$$

$$Kurtosis : K(x) = E\left[\left(\frac{x - E(x)}{\sqrt{Var(x)}}\right)\right]^4 \tag{4}$$

In above equations, mean, *E(x)*, indicates the average of n data points, variance, *Var (x)*, shows how this data is dispersed around the mean, skewness, *S(x)*, measures the expected value, E, of symmetry or asymmetry of the distribution function and kurtosis, *K(x)*, measures the expected value, E, for the height or the degree of peakedness of a data distribution [32].

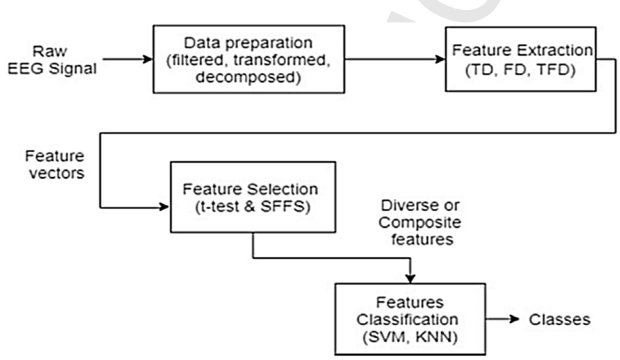

Fig. 2 – Architectual structure of the Proposed EEG signal classifiers.

Then, features were selected according to the following objectives: a) attributes should be a good representative of each signal, b) the extracted attribute must have different values in different classes and, c) the attribute must not be sensitive to the external conditions (e.g., noise, size, angle, etc.) [33].

4.1. *Feature extraction*

In this work, we investigated three different domains of Time, Frequency and Wavelet (time-frequency). Fig. 4 shows the feature extraction process of this work. The figure shows the raw EEG signal was pre-processed in time domain (TD), frequency domain (FD) and time-frequency domains (TFD) using Butterworth filter, Fourier Transform and wavelet transform, respectively. Different waves of Delta, Theta, Alpha, Beta and Gamma defined with their range of frequencies were used in each domain to extract relevant features. Statistical features of mean, variance, skewness and kurtosis was extracted from each brain band. Thus, four features were extracted from each brain wave band and twenty features were extracted in total in each domain. Since, the whole process was performed in all three TD, FD and TFD we ended up with 60 features used for classification. Prior to classification, feature selection was performed to increase system performance. Feature selection strategy used in this study is explained in section IV.2 in detail.

4.1.1. *Time-Domain (TD)*

To extract the four statistical features of mean, variance, skewness and kurtosis, we designed the third order Butterworth bandpass filter and applied that on the entire EEG dataset. The applied filter relates the output signal to the input signal, as shown in the following equations:

$$y(n) = \sum_{i=0}^{N} a_i x(n-i) + \sum_{j=1}^{N} b_j y(n-j) \tag{5}$$

The z-transfer function of the used filter is as follows:

$$H(z) = \frac{\sum_{i=0}^{N} a_i z^{-i}}{1 + \sum_{j=1}^{N} b_j z^{-j}} \tag{6}$$

where *x(n)* is the input signal, *y(n)* is the output signal, N is the filter's order and a_i and b_j are the filter's coefficients [34,35].

To apply the designed filter, we used five brain waves with frequency ranges, Delta (0.4–4 Hz), Theta (4–8 Hz), Alpha (8–12 Hz), Beta (12–30 Hz), Gamma (30–70 Hz), as Low Pass and High Pass cutoff frequencies. As a result, twenty statistical features were extracted from each single channel EEG segment for each of the five different categories (A, B, C, D, E) that we had (Fig. 4).

4.1.2. *Frequency-Domain (FD)*

We continued our investigation by analyzing the signal in the frequency domains. To extract the required features in this

Please cite this article in press as: Savadkoohi M, Oladduni T. A machine learning approach to epileptic seizure prediction using Electroencephalogram (EEG) Signal. Biocybern Biomed Eng (2020), https://doi.org/10.1016/j.bbe.2020.07.004

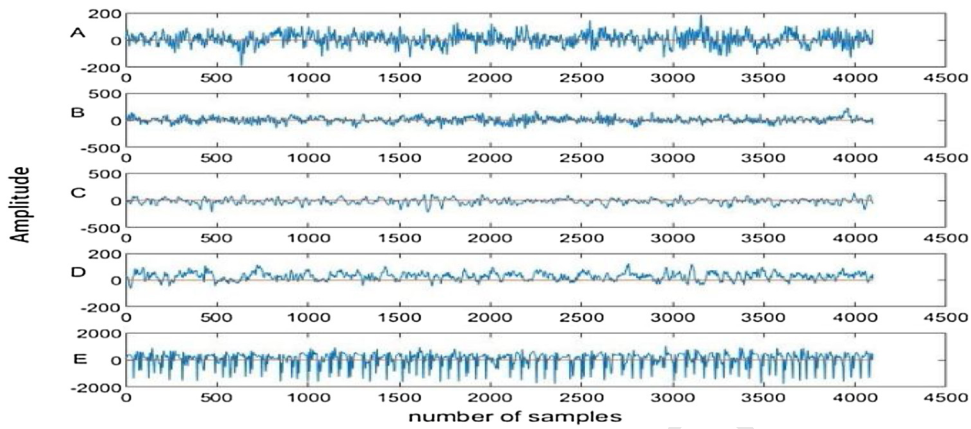

Fig. 3 – Sample EEG signal belonging to subsets A, B, C, D and E.

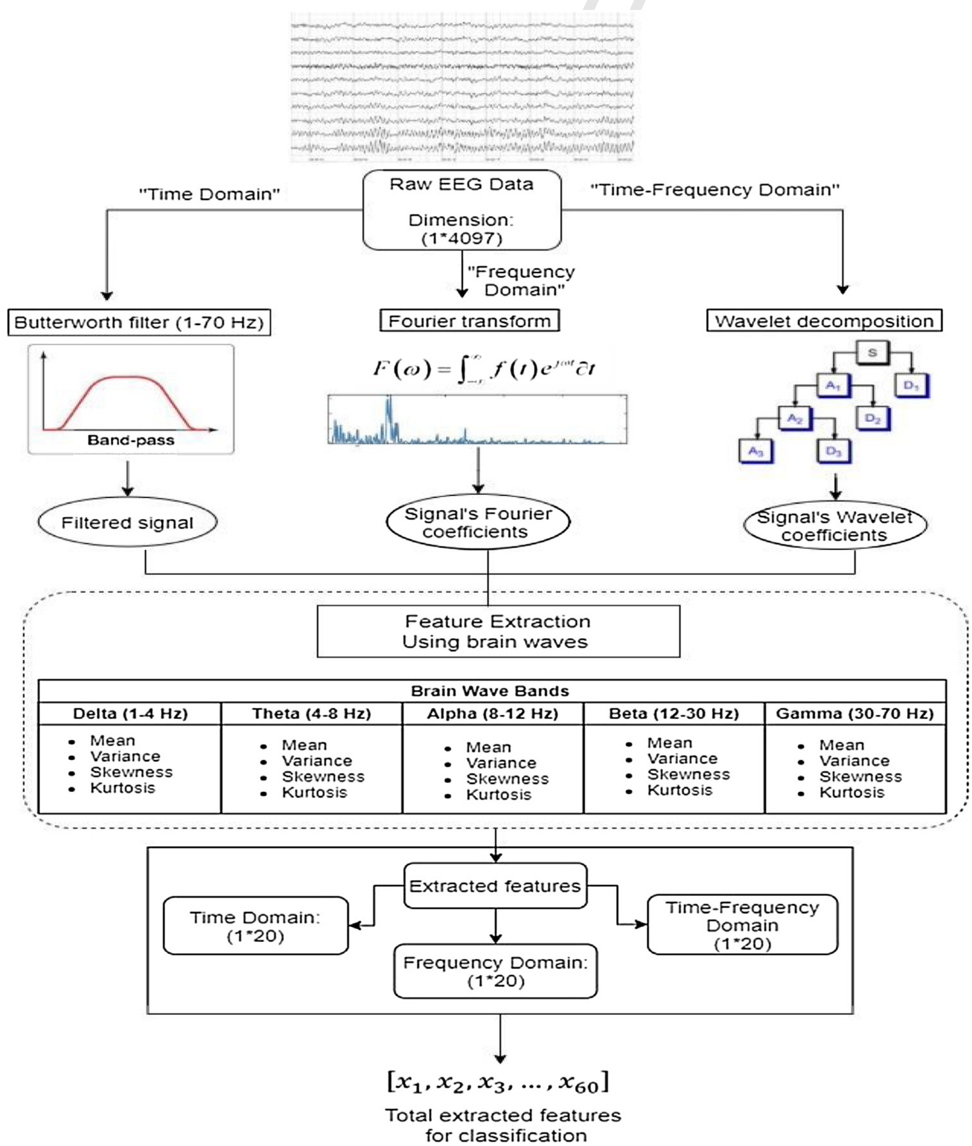

Fig. 4 – Feature Extraction Architecture.

Please cite this article in press as: Savadkoohi M, Oladduni T. A machine learning approach to epileptic seizure prediction using Electroencephalogram (EEG) Signal. Biocybern Biomed Eng (2020), https://doi.org/10.1016/j.bbe.2020.07.004

domain, we used the same five brain waves to specify the frequency bands. A Fourier transformation of the raw EEG signal from time domain to the frequency domain was done. Since the transformed EEG signal is to be divided into the frequency sub-bands of Delta, Theta, Alpha, Beta, Gamma, we created a loop in our code with the number of frequency bands, and in each repetition, we created a new EEG signal and separated the frequency band information in the frequency domain. The statistical properties were then extracted from the amplitude of the frequency coefficients this time (Fig. 4).

According to the Fourier theory, any periodic signal can be decomposed into infinities sum of harmonically related sinusoids and written as the following equation [36,37]:

$$f(x) = a_0 + \sum_{n=1}^{\infty}[a_n Cos(n\omega x) + b_n Sin(n\omega x)] \tag{7}$$

As shown in Eq. (7), $\omega = 2\pi f$ is the angular frequency in which, f is the frequency of the signal and a_o, a_n, b_n are constants called the coefficients of the series.

The Fourier coefficients contain the amplitude and phase of the high order harmonics [36,37]. To extract the statistical features in the frequency domain, we used the amplitude of these harmonics. These coefficients are symmetric in frequency domain. This is due to the fact that any time domain input function for Fourier transform, can be an even $e(x)$, odd $o(x)$ or sum of an even and an odd function $f(x) = e(x) + o(x)$.

Since an even function is symmetric with respect to y axis, the integral of its positive and the negative halves is: $\int_{-\infty}^{+\infty} e(x)dx = 2\int_{0}^{\infty} e(x)dx$. Similarly, the integrals of the positive and negative sides cancel out each other in odd functions, because the symmetry is an inversion at the origin, $\int_{-\infty}^{+\infty} O(x)dx = 0$. On the other hand, the Fourier transform of any function can be written as follows:

$$F(q) = \int_{-\infty}^{+\infty} f(x)e^{-iqx}dx \tag{7}$$

where $f(x)$ denotes the signal in time domain and the $F(q)$ denotes the signal in frequency domain.

In general, the Fourier transform input and output functions are complex functions. Therefore, for a complex function: $f(x) = re(x) + im(x)$, the Fourier transform integral is broken into four components:

$$F(q) = \int_{-\infty}^{+\infty} re(x)Cos(qx)dx - i\int_{-\infty}^{+\infty} re(x)Sin(qx)dx + i\int_{-\infty}^{+\infty} im(x)Cos(qx)dx - i^2\int_{-\infty}^{+\infty} im(x)Sin(qx)dx \tag{8}$$

where: $e^{-iqx} = Cos\,(qx) - iSin(qx)$

The Fourier transform will be symmetric just as the even-odd symmetry does, if either the imaginary or the real part of the input function is zero. As a result, it is not required to calculate all the Fourier coefficients while half of them are zero [36,38]. Due to this symmetric property, we just used half of the Fourier transform coefficients in our code to extract the desired features which leads to saving effort and computing power (Fig. 5).

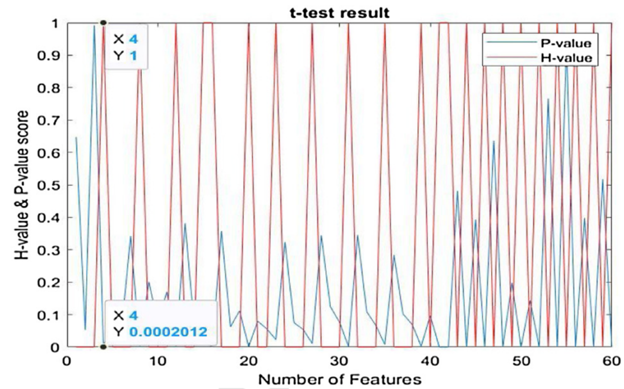

Fig. 5 – T-test result; including features' H-value and P-value scores.

4.1.3. *Time-Frequency Domain (TFD) (wavelet transform)*

One of the drawbacks of Fourier transforms is that it cannot follow instantaneous changes in the signal. To solve this problem, infinite sinusoidal waves or Short-time Fourier transform (STFT) are alternatives. However, this approach is not practically possible due to finite computer memory and frequency resolution limitations [39]. Therefore, wavelet transform (WT) is a better option. It is a method which deconstructs the continuous signal into mini waves (wavelets) that are limited in time and frequency. In other words, wavelet transform has mother-wavelets with variety of frequencies, which makes it possible to follow even the slightest changes in the signal [20,40,41].

To obtain the frequency bands of each wave through the wavelet transform, we used the method explained in [41]. According to the exploited method, the sampling frequency of the EEG signal was changed from 173.61 to 120 Hz initially. Then, a fourth-level discrete wavelet transform was applied to break the EEG into its sub-band components. After this fourth-level decomposition and computing the coefficients of each sub-band, the EEG signal was broken into five frequency levels, which are approximated as delta, theta, alpha, beta, and gamma and the desired statistical features were extracted from the calculated wavelet coefficients.

There are different types of wavelets which differ from each other in terms of shape, smoothness, compactness, etc. Among these various wavelet bases, the Daubechies group is known as an efficient filter and implemented due to its orthogonal properties [41]. For wavelet decomposition of EEG signal in this work, we employed Daubechies order-6 (DB6) wavelet as the basis. A discrete signal can be represented with the following equation in $l^2\,(Z)$,

$$f[n] = \frac{1}{\sqrt{M}}\sum_{k} W_{\Phi}[j_0,k]\Phi_{j_{0,k}}[n] + \frac{1}{\sqrt{M}}\sum_{j=j_0}^{\infty}\sum_{k} W_{\Phi}[j,k]\Psi_{j,k}[n] \quad (9)$$

Where $f[n]$, $\Phi_{j_{0,k}}[n]$, $\Psi_{j,k}[n]$ are discrete functions defined in M points totally and $l^2(Z) = \left\{ f[n] \,|\, \sum_{n=-\infty}^{\infty} |f[n]|^2 < \infty \right\}$. Also, due to the orthogonal property of $\Phi_{j_{0,k}}[n]$ and $\Psi_{j,k}[n]$, the wavelet coefficients can be simply obtained through the inner product rule [39]. As a result, the wavelet coefficients which are known as approximation and details are as follows:

$$W_{\Phi}[j_0,k] = \frac{1}{\sqrt{M}}\sum_{n} f[n]\Phi_{j_{0},k}[n]$$

$$W_{\Psi}[j,k] = \frac{1}{\sqrt{M}}\sum_{n} f[n]\Psi_{j,k}[n](j \geq j_0) \quad (10)$$

The approximation and detail coefficients are the two output coefficients of wavelet transform and represent the output of the low pass filter and high pass filter respectively [39]. The statistical features were calculated from these coefficients.

4.2. *Feature selection*

Extracting attributes is a complex process because some of the extracted features may not be relevant. Therefore, we prioritized the importance of obtaining all necessary and sufficient features for building an efficient classifier. In view of this, we used two criteria to choose our attributes. The first criterion is looking at a single feature and evaluating it in order to see if it is a good representation of the signal and the pattern behind them. The second criterion looks at each feature along with other attributes to find the combination of those ones that have the best performance beside each other [29,30,33]. Using this approach, we argue that redundancy is eliminated, thereby improving the fitness and performance of the proposed classifier.

T-test technique was implemented to find statistically significant information [19]. This method is a scalar method and looks only at one feature at a time and investigates the null and alternative hypotheses. Our null hypothesis assumes there is no statistically significant relationship between a given feature and the population. Therefore, removing that particular feature doesn't affect the model's performance. On the other hand, the alternative hypothesis, emphasizes on the feature's importance and states that there is a statistically significant difference between the given feature and the population. The test for each property produces two parameters, H-value and P-value. If the feature's mean value has a considerable difference in two classes, the alternative hypothesis is valid. This indicates that the feature is important and should be kept. On the other hand, if the mean values don't show a remarkable difference in two classes, the null hypothesis is true and the H-value is set to zero in MatLab code which means that the attribute is not meaningful at all and should be deleted. To compute H-value following equation is used:

$$H = \frac{12}{N(N+1)}\sum_{i=1}^{k} n_i(R_i - R) \quad (11)$$

where n_i is the ith sample of k populations, $N = \sum_{i=1}^{k} n_i$ is the largest sample observation rank, R_i is the sum of the ranks in the ith sample, $R_i = \frac{R_i}{n_i}$ denotes each sample mean and $R = \frac{N+1}{2}$ represents the overall mean. If the difference between sample mean and overall mean $(R_i - R)$ is large, the null hypothesis is rejected. More specifically, null hypothesis can be rejected on significance level α, when:

$$H > \chi^2_{1-\alpha,k-1} \quad (12)$$

where $\chi^2_{1-\alpha,k-1}$ is $(1-\alpha)$-quantile or the critical point of the chi-square distribution with $k-1$ populations. Chi-square distribution is a popular method with different statistical applications. In this case, chi-square distribution evaluates if the variables are truly independent of each other. If so, we can say two samples come from different populations and have meaningful differences [42,43].

The P-value is then used to rank the remained attributes in terms of their significance; the smaller the P-value, the better the rating will be. P-value is typically calculated by writing codes in MatLab, spreadsheet programs or statistical software. However, the computation procedure is the same and follow similar steps. In order to calculate the P-value, we firstly determine Z-score through following formula:

$$Z = \frac{x - \mu}{\sigma/\sqrt{n}} \quad (13)$$

where x is the sample mean, μ is the population mean, σ is standard deviation and n is the sample size. Z-score calculates the number of standard deviations from the mean for particular data points. Having Z-score, we can find the probability (P-value) by which the null hypothesis is rejected. If the P-value is less than significance value or alpha (α), the feature makes a statistical difference and is important [42,43] The P-value threshold (α) in this work was considered as 0.05. Fig. 5 shows the H-value and P-value scores of all sixty features that went through T-test process in our work. As seen in the plot, all those features that are statistically significant $(\alpha < 0.05)$, have been kept by T-test and the corresponding H-value score of 1 has been given to them. To make it clearer, we have highlighted the 4th feature which has a P-value of 0.0002 and H-value score of 1.

According to the T-test results, 22 of the extracted features are statistically significant and can be used for the purpose of classification. More information about the T-test, P-value and H-value can be found here [44].

In series with T-test, for selecting the proper features is Sequential Forward Floating Selection (SFFS) which is a vector method and chooses the best features through a repetitive process [44]. When we are dealing with numerous datapoints, SFFS is a time-consuming method. Therefore, we used T-test initially to remove the non-statistically significant attributes to accelerate the SFFS process. This technique selects the best feature in the first iteration, then looks for another feature that has the best performance along with the previous one and this process continues until the end.

SFFS is in the family of sequential feature selection methods and is used to reduce the initial d-dimensional feature space: $Y = \{y_1, y_2, y_3, \ldots, y_d\}$ to a k-dimensional feature subspace: $X_k = \{x_j | j = 1, 2, 3, \ldots, k;\ x_j \in Y\}$ where $k < d$. Ultimately, the program will choose the k best combination of features that will increase classification performance. Following is an intuitive representation of how SFFS algorithm selects features incrementally to increase performance:

$$X_0 = \{\emptyset\}, k = (0)$$

$$x^+ = argmax\ f(x_k + x);\ where\ (x \in Y - X_k)$$

$$If\ f(x_k + x) > f(x_k)$$

$$X_{k+1} = X_k + x^+$$

$$k++$$

Go to step 3

$$x^- = argmax\ f(x_k - x);\ where\ (x \in X_k)$$

$$If\ f(x_k - x) > f(x_k)$$

$$X_{k-1} = X_k - x^-$$

$$k--$$

Go to step 2

The algorithm initially starts with an empty set so that the size of subset, k, is zero. Step 2 is called inclusion step and adds a new feature, x^+, to the subset X_k if it increases the performance of the model significantly. Effectiveness of new features to make a better model from the subset is evaluated by the criterion function (f). Inversely, step 3 is called exclusion step, and removes a feature, x^- from subset X_k, if its removal makes a better model from the subset and thus increases performance. Steps 2 and 3 are repeated until k equals the number of desired features d and the loop is terminated [45,46].

SFFS graph (Fig. 6) shows that with combination of only five features we can achieve a 100 % representation of the dataset. This implies that these are the necessary and sufficient features representing the dataset for the fitness of the proposed model. As shown in the graph, the performance with this combination is maximum, and the low number of these features reduced the dimensionality of the dataset. The graph also demonstrates that after 11 features, the classifier performance drops to 50% accuracy with additional features. In other words, using extra features leads to system deficiency and reduces performance significantly.

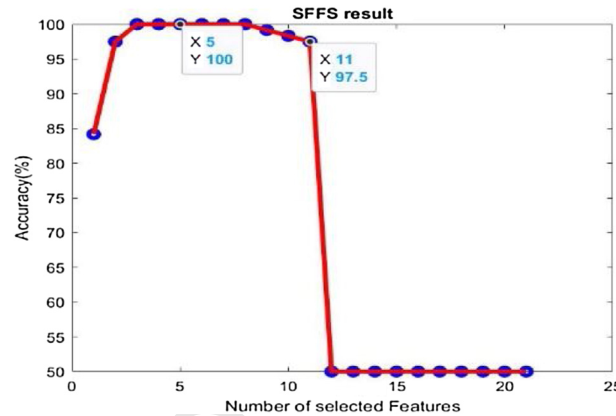

Fig. 6 – Feature selection graph determined by SFFS method.

4.3. *Classification*

We divided our dataset into two for training and testing the learning algorithm through a technique called classification. There are different types of classifiers for categorizing information [47]. In this experiment, we used K-nearest neighbors (KNN) and Support Vector Machine (SVM) learning algorithms to demonstrate the effectiveness of our proposed classification approach. For the KNN, k value was set to three and then the distance between each test sample with all training data was calculated through the Euclidean distance function [48]. With the SVM classifier, we found the best line, hyper plane, which has the longest distance from the nearest data points in both classes [48]. The results of features classification will be explained later in section V.

4.4. *Performance evaluation*

In order to evaluate the result of our analysis and to find out how well the model can generalize; we did evaluation of classifier performances with 5-fold cross validation and assessed its efficiency by calculating accuracy and confusion matrixes [30]. K-fold cross validation method has the advantage of using all instances in a dataset for either training or testing, where each instance is employed for validation exactly once [17,18]. Cross validation reduces overfitting.

5. Result and discussion

Datasets were loaded into the MatLab 2018 programming environment and the code was written to convert the file from. txt to. mat format in order to save data as matrixes. The dataset went through optimal feature extraction before feeding signal into classifiers. We used third order Butterworth bypass filter and applied that on all EEG datasets. We defined low pass and high pass cutoff frequencies based on the brain Delta, Theta, Alpha, Beta and Gamma waves. After this, we used the produced filtered datasets to extract four statistical features from each segment of EEG in the time domain. Since we considered each of the five brain frequency bands, twenty

Please cite this article in press as: Savadkoohi M, Oladduni T. A machine learning approach to epileptic seizure prediction using Electroencephalogram (EEG) Signal. Biocybern Biomed Eng (2020), https://doi.org/10.1016/j.bbe.2020.07.004

Table 1 – Classification results by KNN using time-domain properties.

Total classification Accuracy	Sensivity	Specificity	Confusion Matrix	Class 2 Classification Accuracy	Class 1 Classification Accuracy	Classes
99.5 %	99%	100%	100 0 1 99	100%	99%	A-E
98.5%	97%	100%	100 0 3 97	100%	97%	B-E
98.5%	98%	99%	99 1 2 98	99%	98%	C-E
96.5%	97%	96%	96 4 3 97	96%	97%	D-E
97.2%	98.5%	92.2%	392 8 6 94	98%	94%	ABCD-E
96%	96%	96%	192 8 8 192	96%	96%	AB-CD

features were totally acquired for each single channel recorded EEG sample. Thus, we reduced our huge sample size, to only 20 important features for each segment of our dataset in the time domain.

Then we extracted the EEG waveform information in the frequency domain through Fourier transform which is EEG signal transmission from time domain to frequency domain. Since the frequency coefficients are symmetric in the frequency domain, we therefore employed half of these coefficients. The frequency coefficients are different; we extracted the statistical characteristic from the amplitude of these frequency coefficients. Again, we derived 20 important features in the frequency domain considering all brain range of frequencies. To analyze the signal' time and frequency properties simultaneously, we also applied the wavelet transform on the EEG data and derived the features from the wavelet coefficients. The extracted features in all three domains are mean, variance, skewness and kurtosis.

Extracted features were then tested through the feature selection methods that were explained earlier. Fig. 5 shows the result of the feature selection. According to the derived graph, with the first five features combination (mean, variance, skewness and kurtosis belong to time domain-delta band and mean belongs to the time domain-theta band) the system performance is maximum.

Classification results as well as their performances are shown in Table 1–6. Table 1–2 represent SVM and KNN classification accuracy, specificity, sensitivity and confusion matrix results in the time-domain, tables 3–4 summarize the performance of our used classifiers in the frequency-domain and tables 5–6 demonstrate the classification results of the time-frequency domain using wavelet transform.

Each table is composed of seven columns. We have considered variety of different cases to classify the features extracted from each dataset. The first column of tables represents each group for which the classification was performed as follows:

- A-E: Healthy people (eyes-open) vs. Epileptic patients (during seizure activity)
- B-E: Healthy people (eyes-closed) vs. Epileptic patients (during seizure activity)
- C-E: Seizure-controlled patients (signal of epileptogenic zone) vs. Epileptic patients (during seizure activity)
- D-E: Seizure-controlled patients (signal of opposite brain hemisphere) vs. Patients (during seizure activity)
- ABCD-E: All Healthy and seizure-controlled people vs. Epileptic patients (during seizure activity)
- AB-CD: Healthy people vs. Seizure controlled epileptic patients.

Table 2 – Classification results by SVM using time-domain properties.

Total classification Accuracy	Sensivity	Specificity	Confusion Matrix	Class 2 Classification Accuracy	Class 1 Classification Accuracy	Classes
99.5 %	99%	100%	100 0 1 99	100%	99%	A-E
98.5%	97%	100%	100 0 3 97	100%	97%	B-E
95.5%	95%	96%	96 4 5 95	96%	95%	C-E
96.5%	95.2%	98%	98 2 5 95	98%	95%	D-E
98%	98.7%	95%	395 5 5 95	98.75%	95%	ABCD-E
96.75%	98%	95.6%	191 9 4 196	95.5%	98%	AB-CD

Table 3 – KNN Result using Frequency-domain properties.

Total classification Accuracy	Sensivity	Specificity	Confusion Matrix	Class 2 Classification Accuracy	Class 1 Classification Accuracy	Classes
99 %	98%	100%	100 0 2 98	98%	100%	A-E
98%	96.1%	100%	100 0 4 96	96%	100%	B-E
97.5%	95.2%	100%	100 0 5 95	95.2%	100%	C-E
93.5%	94%	93.1%	93 7 6 94	94%	93%	D-E
97%	98%	93%	393 7 8 92	98%	93%	ABCD-E
88.25%	90%	86.6%	172 28 19 181	90%	86.6%	AB-CD

Table 4 – SVM Result using Frequency- domain properties.

Total classification Accuracy	Sensivity	Specificity	Confusion Matrix	Class 2 Classification Accuracy	Class 1 Classification Accuracy	Classes
100%	100%	100%	100 0 0 100	100%	100%	A-E
100%	100%	100%	100 0 0 100	100%	100%	B-E
100%	100%	100%	100 0 0 100	100%	100%	C-E
100%	100%	100%	100 0 0 100	100%	100%	D-E
100%	100%	100%	400 0 0 100	100%	100%	ABCD-E
95%	95.4%	94.5%	189 11 9 191	95.4%	94.5%	AB-CD

Table 5 – Classification results by KNN using Time-Frequency features combination (wavelet transform).

Total classification Accuracy	Sensivity	Specificity	Confusion Matrix	Class 2 Classification Accuracy	Class 1 Classification Accuracy	Classes
99.5 %	99%	100%	100 0 1 99	100%	99%	A-E
99%	98%	100%	100 0 2 98	100%	98%	B-E
97%	95.2%	99%	99 1 5 95	97%	95%	C-E
93.5%	94%	93%	93 7 6 94	94%	93%	D-E
97.2%	98.2%	93%	393 7 7 93	98.25%	93%	ABCD-E
89%	87.5%	90.6%	182 18 26 174	91%	87%	AB-CD

The next two columns of each table show the classification accuracy of the first and second classes individually, forth column shows the result of confusion matrix calculated for each classification problem and the last three columns show specificity, sensitivity and total classification accuracy respectively.

In the time-domain, the highest accuracy of 99.5% was derived by both the KNN and SVM classifiers for A & E datasets which are related to the healthy people, eyes closed, and epilepsy patients during the seizure activity intervals respectively. The performance of classifiers in the frequency domain, confirmed the previous results. The efficiency of KNN classifier for discrimination of the same classes (A & E) is 99% while SVM shows highest accuracy of 100% for almost all different tested categories (Tables 1–4).

Combination of time and frequency properties or using wavelet transform on the signal, gives almost the same result.

Table 6 – Classification results by SVM using Time-Frequency features combination (wavelet transform).

Total classification Accuracy	Sensivity	Specificity	Confusion Matrix	Class 2 Classification Accuracy	Class 1 Classification Accuracy	Classes
100%	100%	100%	**100 0** **0 100**	100%	100%	A-E
99.5%	99%	100%	**100 0** **1 99**	100%	99%	B-E
98.5%	100%	97.1%	**97 3** **0 100**	97%	100%	C-E
100%	100%	100%	**100 0** **0 100**	100%	100%	D-E
99.8%	100%	99%	**399 1** **0 100**	99.75%	100%	ABCD-E
96%	96%	96%	**192 8** **8 192**	96%	96%	AB-CD

Table 7 – Highest classification accuracy of each classifier at Time-domain, Frequency-domain, and time-frequency (wavelet transform) combination domain.

Classifier	Maximum Accuracy (%)		
	TD	FD	WT
SVM	99.5	100	100
KNN	99.5	99	99.5

Table 8 – Comparison with State-of-the-Art.

Study	Highest determined results (%)			
	Method	Accuracy	Sensitivity	Specificity
Acharya et al. [21]	CNN	88.67	95	90
Guha et al. [22]	DNN	80	80	N/A
Patidar et al. [23]	LS-SVM	97.75	97	99
Martis et al. [50]	Decision Tree	95.33	98	97
Present work	SVM	100	100	100
Present work	KNN	99.5	99	100

According to the numbers achieved in the time-frequency domain the highest accuracy is again related to SVM classifier for A&E (healthy-eyes-closed and epileptic-seizure activity) and D&E (epileptic-seizure-free and epileptic-seizure activity) classes. Tables 7 shows the highest accuracy of each classifier in TD, FD and WT using A and E datasets.

6. Conclusion

In this study we have designed, developed and evaluated an epileptic seizure predictive model. We used the available EEG dataset in [25] recorded for the group of five healthy participants in two eyes-closed and eyes-open situations as well as five epilepsy patients during seizure free and epileptic seizure intervals. Each dataset is composed of 100 EEG segments recorded by the sampling rate of 173.61 Hertz and totally 4097 data-points. In order to reduce the size of datasets, we extracted four statistical features of mean, standard deviation, skewness and kurtosis in the time, frequency and time-frequency domains.

T-test and SFFS, were applied in series to select and determine all necessary and sufficient features to build a reliable classifier. The selected features were fed into the SVM and KNN classifiers and the results showed the highest accuracy of 100% and 99.5% respectively. The results were tested by k-fold cross validation technique. By this method we divided the whole dataset into k subsets. Each subset used k times for training and exactly once for testing the classifiers performance. Consequently, our results show good agreement with the previous studies done on the same dataset.

As explained in section II, many researchers have analyzed EEG signal using different approaches and achieved wide range of accuracies, 76%-100%, to characterize A to E classes [18–26]. For the same classes that we have also considered and analyzed in this work, Bhattacharyya et al. [25] got accuracies between 98-100%, using tunable-Q wavelet transform (TQWT) feature extraction method. Just like this study, they employed SVM classifier and their result is in a good agreement with ours. However, their proposed model is Quality factor (Q) based, application specific and works better for the high frequency scale of the signal. In other words, in order to attain high performance, the optimal selection of Q and redundancy parameter (R) is required and needs to be defined precisely after multiple trials [20]. Providing the same level of accuracy, this work takes advantage of both SVM and KNN classifiers and is not application specific limited. The feature engineering approach in the study doesn't depend on the Quality factor (Q). Furthermore, our work has the capability of robust performance when run on both high and low frequency biological signal.

Also, result from [20] is in strong agreement with our results using the Least-squares support vector machine and fractal dimension of the EEG signal. However, the experiment went through the analytic time frequency flexible wavelet transform (ATFFWT) preprocessing stage. The authors argued that it was necessary in order to accurately detect transient intervals. As shown in our experimental result, we arrived at the same accuracy of result without adding an extra preprocessing stage of ATFFWT, thereby saving computational space and time. Furthermore, compared with the proposed model, the previous work was validated only with one learning

Please cite this article in press as: Savadkoohi M, Oladduni T. A machine learning approach to epileptic seizure prediction using Electroencephalogram (EEG) Signal. Biocybern Biomed Eng (2020), https://doi.org/10.1016/j.bbe.2020.07.004

algorithm, namely, LS-SVM. In contrast the dependability and consistency of the experimental result of the proposed model was validated using two learning algorithms, namely, KNN and SVM. We demonstrated that both KNN and SVM produced competitive results.

The accuracy, specificity and sensitivity performance of both SVM and KNN classifiers confirms the validity of the feature engineering methodology of feature extractions at TD, FD and TFD domains with feature selections using T-test and SFFS of our work. With the highest performance of the SVM classifier's 100% accuracy, sensitivity and specificity and KNN classifier's 99.5%, 99% and 100% accuracy, sensitivity and specificity respectively, our approach showed that 5 features are necessary and sufficient to build a reliable model. Thus, the proposed model is less complicated, reduced training time and efficient. This is in contrast with ReliefF selection method in [49] that required 15 features for achieving optimal classification result with KNN and SVM. Table 8 shows the highest classification accuracy, sensitivity and specificity of the present work. Comparison was made with state-of-the-art. As discussed in section II and seen in Table 8, our results are in good agreement with previous studies and in some cases superior to them in terms of performance.

In view of the above, we argue that the proposed model is reliable and efficient in the detection of epileptic seizures using EEG signal. Furthermore, it is flexible and can be modified easily for different range of frequencies. This advantage makes our approach suitable for other biological data analysis such as: electromyogram (EMG), respiratory and electrocardiogram (ECG) signal.

Implication of the study

In this study we;

1 Designed, developed and evaluated an epileptic seizure classifier
2 EEG signal was pre-processed in time domain (TD), frequency domain (FD) and time-frequency domains (TFD) using Butterworth filter, Fourier Transform and wavelet transform respectively.
3 Different waves of Delta, Theta, Alpha, Beta and Gamma defined with their range of frequencies were used in each domain to extract relevant features.
4 Statistical features of mean, variance, skewness and kurtosis was extracted from each brain band.
5 Combination of T-test and SFFS was used for feature selection and redundancy elimination.
6 SVM and KNN were used as classifiers.

The result of our experiment suggests the following;

1 The feature engineering methodology of feature extractions at TD, FD and TFD domains with feature selections using T-test and SFFS is an effective strategy in building a classifier using EEG signal for epileptic seizure detection.
2 Classification accuracy of 100% and 99.5% is achievable using SVM and KNN classifiers respectively for epileptic seizure detection.
3 Since both classifiers (SVM and KNN) detected the epileptic seizures correctly, therefore, either of them is a reliable epileptic seizure classifier.

Authors statement

We have made all necessary corrections to the paper as recommended by the reviewers. Timothy Oladunni and Marzieh Savadkoohi.

Conflict of interest

None.

REFERENCES

[1] Pannese E. Neurocytology: fine structure of neurons, nerve processes, and neuroglial cells: 2nd fully revised and updated edition. Springer International Publishing; 2015. http://dx.doi.org/10.1007/978-3-319-06856-5
[2] Lent R, Azevedo FAC, Andrade-Moraes CH, Pinto AVO. How many neurons do you have? Some dogmas of quantitative neuroscience under revision. Eur J Neurosci 2012;35:1–9. http://dx.doi.org/10.1111/j.1460-9568.2011.07923.x
[3] Abbott LF, DePasquale B, Memmesheimer RM. Building functional networks of spiking model neurons. Nat Neurosci 2016;19:350–5. http://dx.doi.org/10.1038/nn.4241
[4] Michel CM, Murray MM. Towards the utilization of EEG as a brain imaging tool. Neuroimage 2012;61:371–85. http://dx.doi.org/10.1016/j.neuroimage.2011.12.039
[5] Sharanreddy M, Kulkarni PK. Automated EEG signal analysis for identification of epilepsy seizures and brain tumour. J Med Eng Technol 2013;37:511–9. http://dx.doi.org/10.3109/03091902.2013.837530
[6] Spruston N. Pyramidal neurons: dendritic structure and synaptic integration. Nat Rev Neurosci 2008;9:206–21. http://dx.doi.org/10.1038/nrn2286
[7] Uldry L., Ferrez P.W., Millán J.D.R. (PDF) Feature Selection Methods on Distributed Linear Inverse Solutions for a Non-Invasive Brain-Machine Interface n.d. https://www.researchgate.net/publication/41387051_Feature_Selection_Methods_on_Distributed_Linear_Inverse_Solutions_for_a_Non-Invasive_Brain-Machine_Interface (accessed February 2, 2020).
[8] Adjamian P. The application of electro- and magneto-encephalography in tinnitus research-methods and interpretations. Front Neurol 2014;5. http://dx.doi.org/10.3389/fneur.2014.00228
[9] Strobbe G. Advanced forward models for EEG source imaging; 2014.
[10] Fort GG, Ferri FF. Ferri's clinical advisor 2019: 5 books in 1. 1st ed. Elsevier; 2018.
[11] Adeli H, Zhou Z, Dadmehr N. Analysis of EEG records in an epileptic patient using wavelet transform. J Neurosci Methods 2003;123:69–87. http://dx.doi.org/10.1016/S0165-0270(02)00340-0
[12] Polat K, Günes S. Classification of epileptiform EEG using a hybrid system based on decision tree classifier and fast Fourier transform. Appl Math Comput 2007;187:1017–26. http://dx.doi.org/10.1016/j.amc.2006.09.022

Please cite this article in press as: Savadkoohi M, Oladduni T. A machine learning approach to epileptic seizure prediction using Electroencephalogram (EEG) Signal. Biocybern Biomed Eng (2020), https://doi.org/10.1016/j.bbe.2020.07.004

[13] Srinivasan V, Eswaran C, Sriraam AN. Artificial neural network based epileptic detection using time-domain and frequency-domain features. J Med Syst 2005;29:647–60. http://dx.doi.org/10.1007/s10916-005-6133-1
[14] Übeyli ED, Güler I. Features extracted by eigenvector methods for detecting variability of EEG signals. Pattern Recognit Lett 2007;28:592–603. http://dx.doi.org/10.1016/j.patrec.2006.10.004
[15] Bell MA, Cuevas K. Using EEG to study cognitive development: issues and practices. J Cogn Dev 2012;13: 281–94. http://dx.doi.org/10.1080/15248372.2012.691143
[16] Thatcher RW, Biver CJ, North DM. Z score EEG biofeedback: technical foundations; 2004.
[17] Andrzejak RG, Lehnertz K, Mormann F, Rieke C, David P, Elger CE. Indications of nonlinear deterministic and finite-dimensional structures in time series of brain electrical activity: dependence on recording region and brain state. Physical Review E - Statistical Physics Plasmas Fluids and Related Interdisciplinary Topics 2001;64:8. http://dx.doi.org/10.1103/PhysRevE.64.061907
[18] Epileptologie Bonn / Forschung / AG Lehnertz / EEG Data Download n.d. http://epileptologie-bonn.de/cms/front_content.php?idcat=193&lang=3 (accessed February 3, 2020).
[19] Oweis RJ, Abdulhay EW. Seizure classification in EEG signals utilizing Hilbert-Huang transform. Biomed Eng Online 2011;10:38. http://dx.doi.org/10.1186/1475-925X-10-38
[20] Sharma M, Pachori RB, Rajendra Acharya U. A new approach to characterize epileptic seizures using analytic time-frequency flexible wavelet transform and fractal dimension. Pattern Recognit Lett 2017;94:172–9. http://dx.doi.org/10.1016/j.patrec.2017.03.023
[21] Acharya UR, Oh SL, Hagiwara Y, Tan JH, Adeli H. Deep convolutional neural network for the automated detection and diagnosis of seizure using EEG signals. Comput Biol Med 2018. http://dx.doi.org/10.1016/j.compbiomed.2017.09.017
[22] Guha A, Ghosh S, Roy A. Chatterjee S Epileptic seizure recognition using deep neural network Advances in intelligent systems and computing, vol. 937. Springer Verlag; 2020. p. 21–8. http://dx.doi.org/10.1007/978-981-13-7403-6_3
[23] Patidar S, Panigrahi T. Detection of epileptic seizure using Kraskov entropy applied on tunable-Q wavelet transform of EEG signals. Biomed Signal Process Control 2017. http://dx.doi.org/10.1016/j.bspc.2017.01.001
[24] Zahra A, Kanwal N, ur Rehman N, Ehsan S, McDonald-Maier KD. Seizure detection from EEG signals using multivariate empirical mode decomposition. Comput Biol Med 2017. http://dx.doi.org/10.1016/j.compbiomed.2017.07.010
[25] Bhattacharyya A, Pachori R, Upadhyay A, Acharya U. Tunable-q wavelet transform based multiscale entropy measure for automated classification of epileptic EEG signals. Appl Sci 2017;7:385. http://dx.doi.org/10.3390/app7040385
[26] Richhariya B, Tanveer M. EEG signal classification using universum support vector machine. Expert Syst Appl 2018. http://dx.doi.org/10.1016/j.eswa.2018.03.053
[27] Kaya D. The mRMR-CNN based influential support decision system approach to classify EEG signals. Measurement 2020. http://dx.doi.org/10.1016/j.measurement.2020.107602
[28] Roohi-Azizi M, Azimi L, Heysieattalab S, Aamidfar M. Changes of the brain's bioelectrical activity in cognition, consciousness, and some mental disorders. Med J Islam Repub Iran 2017;31:307–12. http://dx.doi.org/10.14196/mjiri.31.53
[29] Al-Fahoum AS, Al-Fraihat AA. Methods of EEG signal features extraction using linear analysis in frequency and time-frequency domains. ISRN Neurosci 2014;2014:1–7. http://dx.doi.org/10.1155/2014/730218
[30] Amin HU, Mumtaz W, Subhani AR, Saad MNM, Malik AS. Classification of EEG signals based on pattern recognition approach. Front Comput Neurosci 2017;11. http://dx.doi.org/10.3389/fncom.2017.00103
[31] Saini M, Chhikara R. Performance evaluation of DCT and DWT features for blind image steganalysis using neural networks. Int J Comput Appl 2015;114:20–3. http://dx.doi.org/10.5120/19974-1868
[32] Westfall PH. Kurtosis as Peakedness, 1905–2014. R.I.P. Am Stat 2014;68:191–5. http://dx.doi.org/10.1080/00031305.2014.917055
[33] Garrett D, Peterson DA, Anderson CW, Thaut MH. Comparison of linear, nonlinear, and feature selection methods for EEG signal classification. Ieee Trans Neural Syst Rehabil Eng 2003;11. http://dx.doi.org/10.1109/TNSRE.2003.814441
[34] Sloss AN, Dominic Symes, Wright C. ARM system developer's guide : designing and optimizing system software. Elsevier/ Morgan Kaufman; 2004.
[35] Tan L, Jiang J. Infinite impulse response filter design. Digital Signal Processing, Elsevier; 2019. p. 315–419. http://dx.doi.org/10.1016/b978-0-12-815071-9.00008-7
[36] Signals and Systems/Fourier Series - Wikibooks, open books for an open world n.d. https://en.wikibooks.org/wiki/Signals_and_Systems/Fourier_Series (accessed February 3, 2020).
[37] Weisstein EW. Fourier Series – from Wolfram MathWorld n.d. http://mathworld.wolfram.com/FourierSeries.html (accessed February 3, 2020).
[38] Winter R. Symmetry: even, odd, real, imaginary functions :: Advanced Techniques :: Rudi Winter's web space n.d. http://users.aber.ac.uk/ruw/teach/340/ft_symmetry.php (accessed February 3, 2020).
[39] Chun-lin L. A tutorial of the wavelet transform; 2010.
[40] Subasi A. EEG signal classification using wavelet feature extraction and a mixture of expert model. Expert Syst Appl 2007;32:1084–93. http://dx.doi.org/10.1016/j.eswa.2006.02.005
[41] Hsu KC, Yu SN. Detection of seizures in EEG using subband nonlinear parameters and genetic algorithm. Comput Biol Med 2010;40:823–30. http://dx.doi.org/10.1016/j.compbiomed.2010.08.005
[42] Krzywinski M, Altman N. Significance, P values and t-tests. Nat Methods 2013;10:1041–2. http://dx.doi.org/10.1038/nmeth.2698
[43] Ostertagová E, Ostertag O, Kovác J. Methodology and application of the Kruskal-Wallis test. Appl Mech Mater 2014;611:115–20. http://dx.doi.org/10.4028/www.scientific.net/AMM.611.115
[44] Baker MC, Kerr AS, Hames E, Akrofi K. An SFFS technique for EEG feature classification to identify sub-groups. Proceedings - IEEE symposium on Computer-Based Medical Systems; 2012. http://dx.doi.org/10.1109/CBMS.2012.6266361
[45] Peng Y, Wu Z, Jiang J. A novel feature selection approach for biomedical data classification. J Biomed Inform 2010. http://dx.doi.org/10.1016/j.jbi.2009.07.008
[46] Sequential Feature Selector - mlxtend n.d. http://rasbt.github.io/mlxtend/user_guide/feature_selection/SequentialFeatureSelector/#sequential-feature-selector (accessed February 20, 2020).
[47] Hu L-Y, Huang M-W, Ke S-W, Tsai C-F. The distance function effect on k-nearest neighbor classification for medical datasets. SpringerPlus 2016;5:1304. http://dx.doi.org/10.1186/s40064-016-2941-7

[48] Zhang T, Chen W. LMD based features for the automatic seizure detection of EEG signals using SVM. Ieee Trans Neural Syst Rehabil Eng 2017;25:1100–8. http://dx.doi.org/10.1109/TNSRE.2016.2611601
[49] Hernández DE, Trujillo L, Z-Flores E, Villanueva OM, Romo-Fewell O. Detecting epilepsy in EEG signals using time, frequency and time-frequency domain features. Studies in Systems Decision and Control 2018;143:167–82. http://dx.doi.org/10.1007/978-3-319-74060-7_9
[50] Martis RJ, Acharya UR, Tan JH, Petznick A, Yanti R, Chua CK, et al. Application of empirical mode decomposition (EMD) for automated detection of epilepsy using EEG signals. Int J Neural Syst 2012;22. http://dx.doi.org/10.1142/S012906571250027X