ORİJİNAL ARAŞTIRMA ORIGINAL RESEARCH

DOI: 10.5336/biostatic.2018-63886

Classification of EEG Signals for Epileptic Seizures Using Linear and Non-linear Classifiers Based Wavelet Transforms and Information Criteria

Dalgacık Dönüşümleri ve Bilgi Kriterlerini Temel Alan Lineer ve Lineer Olmayan Sınıflandırıcılarla Epileptik Nöbetler İçin EEG Sinyallerinin Sınıflandırılması

Ezgi ÖZER[a],
Ozan KOCADAĞLI[b]

[a]Department of Electrical and Computer Engineering, Nova University of Lisbon, PORTUGAL
[b]Department of Statistics, Mimar Sinan Fine Arts University Faculty of Arts and Sciences, Istanbul, TURKEY

Received: 21.11.2018
Received in revised form: 18.02.2019
Accepted: 19.03.2019
Available Online: 25.03.2019

Correspondence:
Ozan KOCADAĞLI
Mimar Sinan Fine Arts University Faculty of Arts and Sciences, Department of Statistics, İstanbul, TURKEY/TÜRKİYE
ozan.kocadagli@msgsu.edu.tr

ABSTRACT Objective: This study presents an efficient procedure that provides an accurate classification of Electroencephalogram (EEG) signals for the detection of epileptic seizure. Essentially, the proposed procedure hybridizes the linear and nonlinear classifiers with the discrete wavelet transforms (DWT) and principal component analysis (PCA), separately. **Material and Methods:** To classify EEG signals more accurately, the proposed multi-resolution signal processing technique splits them into the detailed partitions with different window-widths, and then decomposes them into detail and approximation coefficients by means of DWT. Thus, many specific latent features that characterize the nonlinear and dynamical structures in the signals can be evaluated from these coefficients. During the model estimation process with multivariate logistic regression (MLR) and artificial neural networks (ANNs), to control the complexity of model and reduce the dimension of feature matrix, PCA is used. In addition, to quantify the complexity and select the best models, the information criteria are considered for both MLR and ANNs. To improve the classification performance, ANNs are trained by various gradient algorithms as well as considering early stopping and cross-validation techniques. **Results:** According to analysis results over the benchmark epilepsy data set released by the Department of Epileptology at University of Bonn, the proposed approach is to bring out 99% accuracy ratios for classifying the epileptic signals. **Conclusion:** This approach not only allows making an efficient analysis of EEG signals for detection of epilepsy, but also provides the best model configurations for ANNs and MLR in terms of reliability and complexity.

Keywords: EEG signal processing; epileptic seizures; discrete wavelet transform; artificial neural networks; multinomial logistic regression; principal component analysis

ÖZET Amaç: Bu çalışma, epileptik nöbetlerin tesbiti için Elektroensefalogram (EEG) sinyallerini doğru sınıflandıran etkin bir yöntem önermektedir. Esas olarak, bu yöntem lineer ve lineer olmayan sınıflandırıcıları, ayrık dalgacık dönüşümleri (ADD) ve temel bileşenler analizi (TBA) ile hibritleştirmektedir. **Gereç ve Yöntemler:** Önerilen çoklu-çözünürlüklü sinyal işleme tekniği, EEG sinyallerinin daha doğru sınıflandırılmak için onları farklı bant genişlikli parçalara bölmekte ve bu parçaları ADD yardımıyla ayrıntı (detail) ve yaklaşım (approximate) katsayılarına ayrıştırmaktadır. Böylece, sinyallerin barındırdığı dinamik ve lineer olmayan yapıları karaterize eden birçok gizli özellik, bu katsayılar üzerinden belirlenmektedir. Çokterimli Logistik Regregresyon (ÇLR) ve Yapay Sinir Ağlarıyla (YSA) model kestirim sürecinde, karmaşıklığı kontrol etmek ve veri matrisini indirgemek için TBA kullanılmıştır. Bunun yanısıra, model karmaşıklığının nicelendirilmesi ve en iyi modellerin belirlenmesi için bilgi kriterlerinden yararlanılmıştır. Doğru sınıflandırma performasını arttırmak için YSA'lar erken durdurma ve çapraz geçerlilik teknikleriyle beraber çeşitli gradyan-tabanlı öğrenme algoritmalarıyla eğitilmiştir. **Bulgular:** Bonn Üniversitesi Epileptoloji bölümünce herkesin kullanımına açılmış epilepsi veri seti üzerinden elde edilen analiz sonuçlarına göre, önerilen yaklaşım epileptik sinyallerin ayrıştırılmasında %99'lara varan doğruluk oranları vermektedir. **Sonuç:** Bu yaklaşım epilepsinin teşhisi için EEG sinyallerinin etkin bir analizini yapmakla kalmayıp, model güvenilirliği ve karmaşıklığı bakımından da ÇLR ve YSA'lar için en iyi model konfigürasyonlarını sağlamaktadır.

Anahtar Kelimeler: EEG sinyal işleme; epileptik sinyal; ayrık dalgacık dönüşümü; yapay sinir ağları; çokterimli lojistik regresyon; temel bileşenler analizi

Epilepsy is one of the most common central nervous disorders, which is temporary abnormal electric discharges in the nerve cells and occur at unpredictable times and usually without warning.[1,2,3] Epileptic seizures trigger the involuntary spasms bringing about serious physical injuries or even death. Generally, to examine the characteristics of seizures and reason of epilepsy, there are various tools such as Computed Tomography (CT), Magnetic Resonance Imaging (MRI) scans, Electroencephalography (EEG) and brain imaging tests as well as the medical history and blood tests. Depending on the case, these tools can be used exclusively or together. Specifically, EEG provides valuable information with helping multi-electrode recording of time-varying the electoral activity in the brain. Thus, saving the functional and instant changes in the brain provides valuable information about the electrical activity of the brain and possible types of seizures.[4] Besides, the epileptic EEG recordings having seizure have a large number of spike signals compared to the normal EEG recordings, and the amplitude increases noticeably.[5]

Other than visiual test, to detect the epileptic seizures by using automated computer software, EEG signals related to brain's electrical activity in real-time are partitioned into the seizure and non-seizure periods by means of unsupervised and supervised techniques. Technically, some shortcomings feced in this process on EEG signals can be handled by the supervised techniques in the discriminative structure.[3,6,7] Regarding to EEG test, another issue is including noisy signals. Besides, some significant information related to the past or future of epileptic seizures can be overlooked through EEG signals, because they might be masked by other biological signals. Hence, these shortcomings are eliminated by signal processing techniques that magnify the relevant information and to extract the latent features from EEG signals.[8] By means of the signal processing techniques such as Lyapunov exponent, Fourier, Hilbert, Wavelet transforms (WTs), etc., the signals are decomposed into the transformed sub-series. Basically, this process is known as multi-resolution signal processing or sub-band coding.

WTs are very efficient tool to extract the explanatory features from brain signals without not so much loss of information and establishes the decomposition from them into the sub-transformed series with different scales at the desired decomposition levels in the specific time interval. In this way, some specific latent features which characterize the dynamical and nonlinear frameworks in the signals can be obtained from these sub-transformed bands.[9-14] In the literature, there are remarkable approaches in which different feature extraction techniques are processed to extract different types of features such as the statistical moments, entropy and metric measures.[15-18]

In the context of seizure detection, the most commonly used techniques are Discriminant Analysis, Logistic Regression, Gaussian Mixture Models, Regression Trees, Random Forest, k-Nearest Neighbor, Naive Bayes Classifier, Kernel Methods, Support Vector Machines (SVMs), ANNs and ANFIS classifiers. According to analysis results reached from current studies in the litereature, hybrid AI techniques give pretty much superior performance to the classical statistical ones, because they estimate much more efficient models in the nonlinear and high dimensional cases. In these studies, WTs are often used together with various classifiers. Especially, ANN classifiers are often used together with WTs due to their flexible and adjustable structures.[10,11,14,19-27] Also, there are many studies in which different feature extraction techniques are perefered with the other classifiers rather than ANNs.[1,12,13,15-17,28-40]

MOTİVATİON AND OVERVİEW

In the context of epilepsy, recently the automated multi-resolution tools are used together with the classical statistical methodologies or AI techniques against the classical visual test. To do this, EEG recordings are initlally partitioned into the sub-series with the predetermined window-widths. After that, these sub-series are decomposed into the transformed sub-series by means of signal processing techniques. Actually, this process brings out delta, theta, alpha, beta and gamma waves having different frequency bands from brain signals.[4,41-43] Generally, to assign a suitable window-width, there are two approaches in the litrrature: the fixed and non-fixed ones. In this study, the automated multi-resolution approach starts a fixed window-width, and then shortens it gradually each iteration until the feature extraction process produces the desired outputs.

To ivestigate the epileptic signals in a short time, discrete wavelet transforms (DWT) are able to establish a successful multi-resolution analysis.[27,44,45] However, the selection of mother wavelets has an important role to find out the charateristics of epileptic behaviours. On the performance of wavelet families with regarding to making an efficient multi-resolution; Amorim et al., Faust et al., and Chen et al. made remarkable studies.[27,46,47] In this study, to imrove the performance of multi-resolution process, various mother wavelets were tested with respect to the referenced studies.

In the proposed multi-resolution process, the statistical indicators obtained from the wavelet coefficients are used as inputs (features) in the estimation procedure of ANN and MLR models. The dimension of data set (training set) changes with the decomposition level of DWTs and number of features. For this reason, the feature matrix should be reduced to control the complexity of estimated models. However, the researchers are not interested so much the complexity of model during estimation process, even if it is directly related to generalization.

In MLR analysis, the excessive linear correlation between features is known as multicollinearity which is a hypothetical decay. To cope with this problem, PCA produces new orthogonal factors called as the principal components. Thus, the dimension of feature matrix can be reduced with respect to the variance explanation percentage of selected components in addition to overcoming the multicollinearity.

After reducing the dimesion of data set, EEG signals are classified with respect to their predetermined classes and the reduced feature set. In this study, to improve the classification performance, ANNs are preferred as the nonlinear classifiers, since they do not need any restrictive assumption and provide very flexible estimation procedure. From the previous researches in the literature, it's well known that they produce superior performance to the linear classifiers such as MLR and discriminant analysis in the high dimensional and nonlinear enviroments. Despite of their advantages, they suffer from some challenges such as the network desing, model complexity, memory allocation and tuning parameters.[48-52]

Another problem related to ANNs arises at selection of risk function. For instance, the analysts often perefer the mean squared error (MSE) as a risk function. However, this risk function causes two types errors: approximation and estimation errors where they are directly related to the complexity of model and must be reduced simultaneously.[52-55] In the proposed procedure, to handle the complexity problem during the training process of ANNs, some information criteria are used such as Akaike Information Criterion (AIC), Corrected AIC (AICc) and Bayesian Information Criterion (BIC) as well as an early stopping and cross-validation. Also, in this proposed procedure, information criteria allow determing the most efficient number of neurons in the hidden layers of ANNs. To investigate the best training algorithms for ANNs, they are trained various gradient based algorithms such as Gradient Descent (GD) and GD with momentum (GDM), Quasi-Newton known as Broyden, Fletcher, Goldfarb, and Shanno (BFGS), Scaled Conjugate Gradient (SCG) and Levenberg-Marquardt (L-M). To improve the performance of training algorithms, some useful techniques can be considered in terms of tuning parameters and stopping criteria.[48-50,56,57]

Unlike ANNs, MLR doesn't need a complicated methodology in terms of constructing its functional structure and estimation procedure. In addition, MLR doesn't need the multivariate normality assumption different from the discriminant analysis and its derivatives. Generally, the parameters of MLR model are estimated by the maximum likelihood estimation (MLE). As is well-known from the statistical theory, MLE gives better performance in the case of large data.[58,59] But, in the case of limited data, this may turn into a handicap as well. To handle the model complexity and prevent multicollinearity between the independent variables in MLR, PCA and various variable selection procedures such as stepwise, backward, forward, etc. can be used together.[60-62] In this study, to estimate the efficient MLR models and overcome the mentioned problems, firstly the feature matrix is reduced by PCA, and then the best models are determined with respect to the accurate classification ratios and information criteria.

In this study, the main purpose is to propose an efficient procedure allowing the medical experts to make a comprehensive analysis of EEG signals for epilepsy detection using ANNs and MLR, and develop its software for the clinical researches. To introduce the proposed approach, this paper is structured as following. Section 2 gives the main frameworks of DWTs, PCA, MLR and ANNs. Section 3 includes an application in which the proposed and traditional approaches are compared each other over a benchmark EEG data set in the context of detection of epileptic seizures. Section 4 is placed for the analysis results and their interpretions. Lastly, the conclusions and future researches are discussed in Section 5.

METHODOLOGY

The flowchart of the proposed procedure is shown in Figure 1. Initially, EEG signals are received from different electrodes and then they are decomposed into the subseries by DWT. Thus, the original signals are represented in another space where the latent characteristics of them are exposed better. The second stage is to evaluate some significant features over the predetermined window-widths of decomposed subseries. This stage produces a feature matrix that will be used in the classification of EEG signals in the context of epilepsy detection. After that, to control the model complexity and prevent the multicollinearity problem between the feature vectors, the dimension of feature matrix is reduced. Essentially, the multicollinearity is a hypothetical decay for the linear classifiers such as the logistic regression. After reducing the dimension of feature matrix and feature selection, the last stage is to estimate the classification models by means of the linear and nonlinear classifiers: MLR and ANNs, respectively. Before the estimation process, the feature matrix is subject to the cross-validation to ensure more general models. Lastly, to control the model complexity and select the best model configuration, the information criteria are considered as well as the accuracy ratios and error function in the terms of classifying EEG signals correctly.

FEATURE EXTRACTION USING DISCRETE WAVELET TRANSFORM

Wavelet transform is to analyze time series signals in the nonlinear structure at different frequencies and decomposes them by shifting the wavelet in the time axis and changing the size.[63,64] Basically, in the continuous time domain, a wavelet transform of signal x[n] can be defined as

$$wt(s,\tau) = \frac{1}{\sqrt{s}} \int_{-\infty}^{\infty} x(t)\psi^{*}\left(\frac{t-\tau}{s}\right) dt \tag{1}$$

where the function $\psi*(\cdot)$ is a complex conjugate of the shifted and scaled wavelet function $\psi(\cdot)$ Here, the parameter τ shifts wavelet fuction $\psi(\cdot)$ along the time domain. This process is known as "translation". Be-

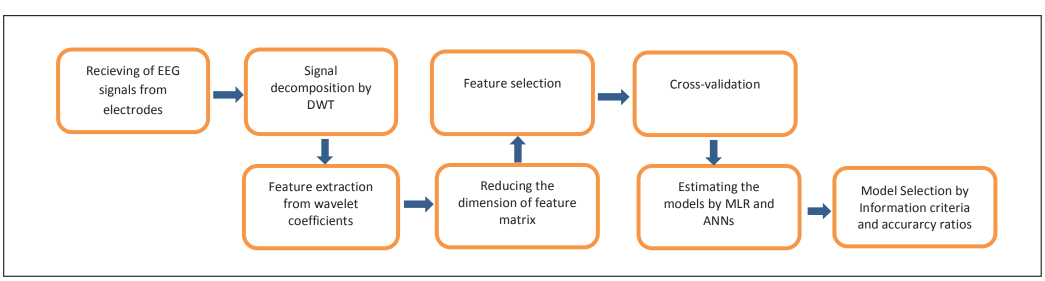

FIGURE 1: The flowchart of the proposed procedure.

sides, the scale parameter s is interested in the stretching the wavelet function $\psi(\cdot)$ This process is called as "dilation". In the context of wavelet transform, a family of wavelet function with scaled and shifted parameters is described as [44,65,66]

$$wt(s,\tau) = \frac{1}{\sqrt{s}} x(t)\psi\left(\frac{t-\tau}{s}\right), \qquad s > 0,\ \tau \in R \tag{2}$$

Generally, the time-frequency analysis is to bring out the masked characteristics in the signal x[n] by means of the decomposition processes. In this framework, both s and τ are changed continuously; however this structure causes the redundant information as well as excessive memory allocation. For this reason, this shortcoming can be handled by discretizing the parameters s and τ instead of continuous domain.

In DWT framework, a set of transformed sub-bands is obtained by using the low-pass h[n] and high-pass g[n] filters where they are known as the wavelet coefficients. By the filtering processes, a signal x(t) can be decomposed into the low and high frequency sub-bands as follows:[67]

$$a_{j,k} = \langle x(t), \phi_{j,k}(t) \rangle = \sum_m h(m-2k) a_{j-1,k} \tag{3}$$

and,

$$d_{j,k} = \langle x(t), \psi_{j,k}(t) \rangle = \sum_m g(m-2k) d_{j-1,k} \tag{4}$$

where $\langle \cdot \rangle$ is the inner product operator. In the high and low frequency components, $a_{j,k} = \langle x(t), \phi_{j,k}(t) \rangle$ and $d_{j,k} = \langle x(t), \psi_{j,k}(t) \rangle$ are called as the detailed and approximate coefficients, respectively. These components at the wavelet decomposition level j are obtained by convolving the approximate and detailed coefficients at decomposition level (j-1) by means of the lowh[n] and highg[n] pass filters. The scheme is exhibited in Figure 2.[44,45,67-70]

In the signal processing by DWT, the choice of an adequate wavelet plays a crucial role to explore main characteristics of the related signal. In the literature of DWT, there exists common wavelet families such as biorthogonal, Coiflets, Daubechies, Discrete Meyer, Haar, Biorthogonal and Symlets.[71-73] To determine a suitable wavelet in the analysis, the researchers are mostly interested in the structures of EEG signals with respect to the related case. There exists remarkable studies for the feature extraction from EEG signals in the literature. Also, the expert knowledge might be enough to use correct wavelet structure in the clinical researches.[27,46,47,66]

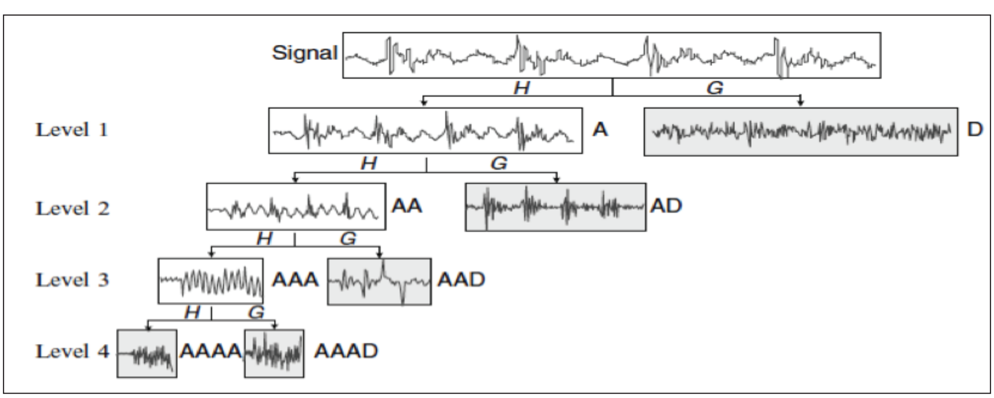

FIGURE 2: G and H are high and low pass filters, respectively.[44,66]

REDUCING THE DIMENSION OF FEATURE MATRIX

Principal component analysis is a statistical technique that can be used to reduce the dimension of data set and brings out a new orthogonal structures.[74-76] This new orthogonal set, called as principal components z_i (PCs), essentially is a linear combination of p original variables.[75]

Let $X = (x_{ij})_{n*p}$ be a data set of features $X_1, X_2, \ldots, X_P$ with n observations. An eigenvalue decomposition can be established as $Se_i = \lambda_i e_i \ (i = 1,2, \ldots p)$ where e_i is an eigenvector and $S = \frac{1}{n-1}X'X$ is a covariance matrix. From above equalities, it can be arranged that $(S - \lambda I_P)e = 0$ and $det(S - \lambda I_P) = 0$.[59] If the eigenvalues are ordered from large to small such as $\lambda_1 \geqslant \cdots \geqslant \lambda_p \geqslant 0$ and m eigenvectors with the largest eigenvalues $(m \leq p)$ are chosen among them, then most of the variance in the original data can be explained as follows:[58,77,78]

$$
\begin{aligned}
z_1 &= a_{11}x_1 + a_{12}x_2 + \cdots + a_{1p}x_p = a_1^T x \\
z_2 &= a_{21}x_1 + a_{22}x_2 + \cdots + a_{2p}x_p = a_2^T x \\
&\vdots \\
z_m &= a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mp}x_p = a_m^T x
\end{aligned}
\tag{5}
$$

where $z_1, \ldots, z_m$ are the m principal components and a_{ij} is the weight of the variable jth for principal component ith. From Eq. 5, z_1 is the first PC that explains the maximum variance in the model. Here, z_2 is the second PC having the second largest variance, and so on. Thus, PCA not only provides reducing the dimension of data matrix, but also ensures an orthogonal system where the linear correlations between variables are vanished. For this reason, PCA is widely used in the image and signal processing, classification and pattern recognition problems.[79]

MULTIVARIATE LOGISTIC REGRESSION

MLR is a statistical method that can be used for predictive analysis. Similar to the multiple linear regression analysis, it has the dependent Y and independent variables $x = (x_1, \ldots, x_p)$ where dependent one includes the categorical outcomes unlike continuous ones such as $Y = (0,1, \ldots, c-1)$.[80,81] MLR can be expressed by a logit function as follows:[80]

$$
\begin{aligned}
g_j(x) &= ln\left[\frac{P(Y=j\backslash x)}{P(Y=0\backslash x)}\right], \ j = 0, \ldots, c-1 \\
&= \beta_{j0} + \beta_{j1}x_1 + \cdots + \beta_{jp}x_p
\end{aligned}
\tag{6}
$$

where, β_{j0} is the intercept and $\beta_{ji}(i = 1, \ldots, p)$ is unknown parameter. Here, the conditional probability of Y is obtained as:

$$
P\big((Y = j|x)\big) = \frac{e^{g_j(x)}}{\sum_{k=0}^{c-1} e^{g_k(x)}}, \quad j = 0, \ldots, c-1 \tag{7}
$$

To estimate the model parameters in Eq. 7, the maximum likelihood estimation (MLE) can be used. Generally, MLE intends to maximize the probabilities of occurrence of the outcomes. In addition, this method needs the large data set to ensure more accurate parameter estimations.[58,76]

ARCHITECTURE OF ANN CLASSIFIERS BASED MSE

In this study, to estimate the robust classification models, ANNs with different configurations are used. Figure 3 shows a simple structure of ANNs. In this structure, the first layer is directly connected with inputs

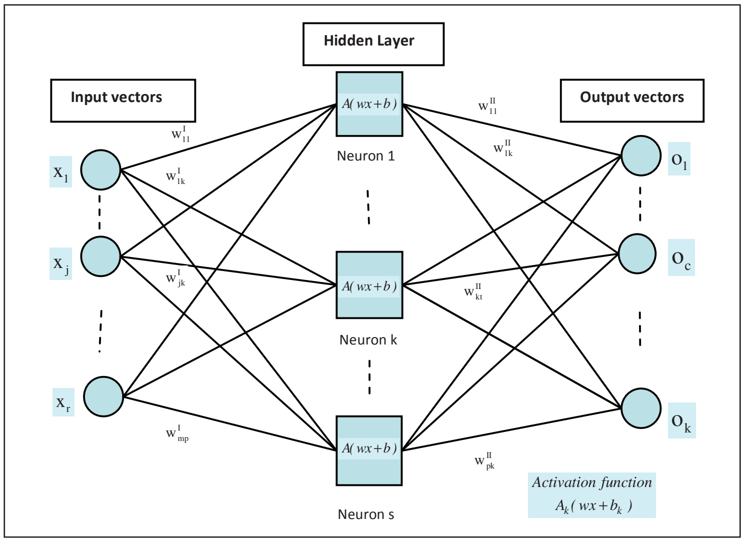

FIGURE 3: The network with one hidden layer.[66]

and the hidden layer includes p number of neurons. The last layer consists of output vector yc that includes totally k number of classes $y_c = [O_1, O_2, \ldots, O_k]$ where c indicates the indices of classes (c=1,2,…,k). In this framework, output yc is related to any one of k classes:

$$y_c \epsilon [\underbrace{\{1,0,0,,\ldots,0\}}_{\text{Class 1}}, \underbrace{\{0,1,0,\ldots,0\}}_{\text{Class 2}}, \{0,0,1,\ldots,0\}, \ldots, \underbrace{\{0,0,0,\ldots,1\}}_{\text{Class k}}] \tag{8}$$

Here, any element of y_c is described as follow:

$$o_c = f(W^I, W^{II}, x) = \frac{e^{\left[b_c^{II} + w_c^{II} A\left(W^I x + b^I\right)\right]}}{\sum_{j=1}^{k} e^{\left[b_j^{II} + w_j^{II} A\left(W^I x + b^I\right)\right]}} \in [0,1] \qquad c = 1,2,\ldots,k \tag{9}$$

Apperantly, Eq. (9) is a soft-max function where

W^I : The weight matrix defined from inputs to the neurons in the hidden layer

W^{II} : The weight matrix defined between the hidden and output layers

x: The input vector includes the features

b^I : The bias vector defined as $b^I = [b_1^I\ b_2^I \ldots b_s^I]$ in the hidden layer

b^{II} : The bias vector defined as $b^{II} = [b_1^{II}\ b_2^{II} \ldots b_k^{II}]$ in the output layer

From definitions above, W^I can be written as

$$W^I = [w_1^I\ w_2^I\ \ldots w_i^I\ \ldots\ w_s^I] \tag{10}$$

s: the number of neurons in the hidden layer

or

$$W^I = \begin{pmatrix} w_{1,1} & w_{1,2} & \cdots & w_{1,r} \\ \vdots & & \ddots & \vdots \\ w_{s,1} & w_{s,2} & \cdots & w_{s,r} \end{pmatrix} \tag{11}$$

where w_i corresponds to a vector denoted as $w_i^I = [w_{i,1}\ w_{i,2}\ \ldots\ w_{i,r}]$ (i = 1, 2, ..., s). Here, w_i includes all the weights among the neuron i and all the inputs. Similarly, W^{II} can be written as:

$$W^{II} = [w_1^{II}\ w_2^{II}\ \dots w_c^{II}\ \dots\ w_k^{II}] \tag{12}$$

k: the total number of classes

$$W^{II} = \begin{pmatrix} w_{1,1} & w_{1,2} & \cdots & w_{1,s} \\ \vdots & & \ddots & \vdots \\ w_{k,1} & w_{k,2} & \cdots & w_{k,s} \end{pmatrix} \tag{13}$$

In Eq. 9, $A(w^I x + b^I): R^s \rightarrow R^s$ corresponds to a vector function that covers s number of activation functions as follow:

$$A(W^I x + b^I) = G\left(A_1(w_1^I x + b_1^I), A_2(w_2^I x + b_2^I), \dots, A_s(w_s^I x + b_s^I)\right) \tag{14}$$

where A_j denotes the j^{th} tangent hyperbolic function in the hidden layer:

$$A_j = \frac{e^{NET_j} + e^{-NET_j}}{e^{NET_j} - e^{-NET_j}};\ NET_j = w_j x + b_j \qquad j = 1, 2, \dots, s. \tag{15}$$

In this framework, the main purpose of ANNs is to minimize the classification error. To do this, the most practical way is to train ANNs with MSE by using the gradient-based algorithms. Essentially, this approach corresponds to minimize L^2 norm in the metric space. Here, as a risk function, MSE can be defined as

$$MSE = \frac{\sum_{i=1}^{N} \sum_{c=1}^{k} (o_{c,i} - y_{c,i})^2}{N \times k} \qquad c = 1,2,\dots,k;\ i = 1, 2, \dots, N.$$

After the training procedure, the last step is to assign the output $y_c = [o_1, o_2, \dots, o_k]$ to any of k classes with respect to Eq. 9. Thus; the components of y_c are transformed into the binary numbers as follow:

$$\chi_{o_c} = \begin{cases} 1, & \text{if } o_c = \max\ [o_1, o_2, \dots, o_k] \\ 0, & \text{Otherwise.} \end{cases} \tag{16}$$

APPLICATION

DATA COLLECTION

In the application, it is used a benchmark data set realesed by Department of Epileptology, University of Bonn is preferred, because it allows comparing the proposed approach with the current ones. Basically, this data set consists of five different groups (A–E) where each one contains 100 single channels of EEG signals. Specifically, each channel covers 4096 samples where they were recorded at 23.6 s duration. Here, groups A and B includes the samples of EEG recordings from five healthy volunteers with eyes open and closed, respectively where they were recorded extracranially by means of a standardized 10-20 electrode placement system. Groups C, D, and E were recorded intracranially from five patients where they were selected from EEG archive of presurgical diagnosis. The samples of EEG signals in set D were recorded within the epileptogenic zone, and those in set C from the hippocampal formation of the opposite hemisphere of the brain. Groups C and D consist of only activity recorded during seizure free intervals whereas grpup E only seizure activity.[11,27]

ANALYSIS

To investigate the epileptic behaviors in the five group sets, the first step is to extract some significant features from EEG signals. Therefore, the proposed procedure partitions each channel into the sub-series with respect to the predetermined fixed window-width. In each iteration, the sizes of window-widths are changed whether the sufficient patterns are provided from EEG signals. Thus, the automated multi-resolution technique continues to look for the latent characteristics in these sub-series with respect to various window-widths and resolution levels until the classifiers produce the desired classification performance. To figure out the latent behaviours of epileptic seizures, some important feautures can be evaluated from the detail and approximation coefficients. To do this, totatally nine feautures are taken into account where they correspond to mean, median, standard deviation, kurtosis, skewness, maximum, minimum, entropy and energy values of the detail and approximation coefficients.

The performance of classifiers depends on the information provided from the detail and approximation coefficients, so various mother wavelets are tried in the processes of DWT. For this reason, the previous studies in the literature are considered as well as trial and error. In the decomposition process of EEG signals, Daubechies 10 (db10) wavelets were preferred, because they provided a sufficient performance for classifiers.

The feature extraction process can be summarized as following. As seen in Figure 4, the series $z_1(t)$ with 4096 observations includes EEG signals from Channel 1 in Group A. Here, $z_1(t)$ is pertitioned into 8 sub-series with 512 samples ($s_i(t)$, i = 1, 2, ….,8). By doing this, more detailed information can be obtained from these small sub-series. For this reason, this process was applied to EEG series in the all the channels. Then, all the partitions $s_i(t)$ were subject to the time frequency decomposition using DWT, separately. As seen in Figure 5, any $s_i(t)$ can be decomposed into the six detail (D_j, j = 1, 2, …,6) and approximation (A_j, j = 1, 2, …,6) coefficients at the sixth level. In this process, the sub-series A_i and D_j are filtered by low (H) and high (G) pass filters, respectively. When the features are being evaluated over these coefficients, the last approximation coefficient at the sixt level and all the details coefficients were considered, because this approach brought out enough information about epileptic seizures.

As seen in Table 1, a feature vector of any partition $s_i(t)$ can be produced by using features calculated from six details and one approximation series. Thus, the feature vector related to $s_i(t)$ consists of 63 statistical indicators (9 features × 7 coefficients). If this process is applied to the other sub-series of $z_1(t)$, then totally 9 feature vectors will be produced.

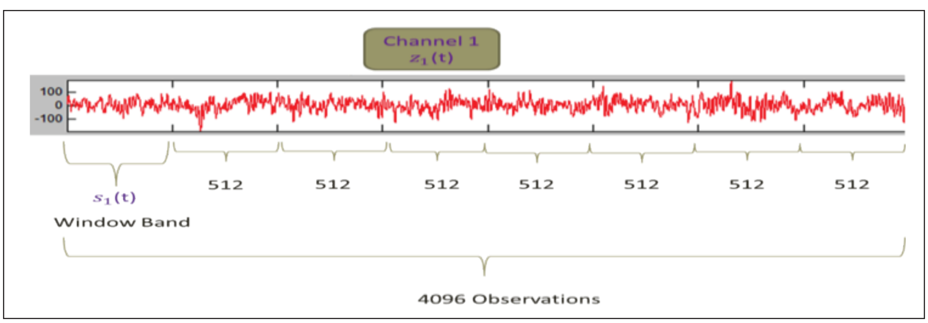

FIGURE 4: Seperation of segment z(t).

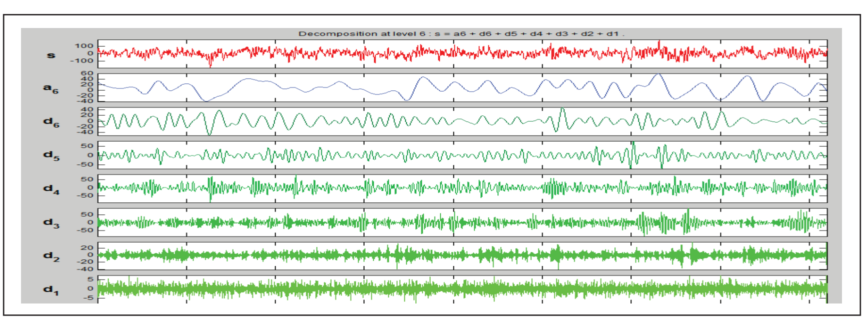

FIGURE 5: Decomposition of s(t) at the sixth level.

TABLE 1: Feature vector of s(t) with 63 components.

d1	d2	d3	d4	d5	d6	a6
Maximum (d1)	Maximum (d2)	Maximum (d3)	Maximum (d4)	Maximum (d5)	Maximum (d6)	Maximum (a6)
Maximum (a6)	Minimum (d2)	Minimum (d3)	Minimum (d4)	Minimum (d5)	Minimum (d6)	Minimum (a6)
Median (d1)	Median (d2)	Median (d3)	Median (d4)	Median (d5)	Median (d6)	Median (a6)
Mean (d1)	Mean (d2)	Mean (d3)	Mean (d4)	Mean (d5)	Mean (d6)	Mean (a6)
Entropy (d1)	Entropy (d2)	Entropy (d3)	Entropy (d4)	Entropy (d5)	Entropy (d6)	Entropy (a6)
Energy (d1)	Energy (d2)	Energy (d3)	Energy (d4)	Energy (d5)	Energy (d6)	Energy (a6)
Skewness (d1)	Skewness (d2)	Skewness (d3)	Skewness (d4)	Skewness (d5)	Skewness (d6)	Skewness (a6)
S.deviation (d1)	S.deviation (d2)	S.deviation (d3)	S.deviation (d4)	S.deviation (d5)	S.deviation (d6)	S.deviation (a6)

In this framework, totally 800 feature vectors can be obtained from 100 Channels in Group A. Similary, for all the groups (A, B, C, D and E), totally 4000 (5×800) feature vectors will be produced. Lastly, at the end of the automated multi-resolution decomposition using the window-widths with 512 samples, a feature matrix with 4000×63 will be created as seen in Table 2. In the context of the supervised learning, the classes corresponding to feature vectors can be designed by using the binary coding with respect to Eq.16 for five groups as following:{0, 0, 0, 0, 1}, {0, 0, 0, 1, 0}, {0, 0, 1, 0, 0}, {0, 1, 0, 0, 0} and {1, 0, 0, 0,0}.[66]

In the estimation procedure, to control the complexity of model, the excessive number of features is eliminated by PCA. Thus, the elimination of some features allows to estimate more robust models by reducing the dimension of feature matrix in addition to reducing unnecessary memory usage. After this process, the feature matrix is used to estimate robust models by MLR and ANNs. To improve the classification performance, ANNs are trained by the gradient-based algorithms. In Figure 6, the scheme of the proposed approach is given in detail.

TABLE 2: Training data set obtained from DWT for the window-widths with 512 samples at the sixth level.[66]

9 Features from D1	9 Features from D2	. . .	9 Features from D5	9 Features from D6	Yc
					1 0 0 0 0
		A 8x100 = 800 observations			. . .
					0 1 0 0 0
		B 8x100 = 800 observations			. . .
					0 0 1 0 0
		C 8x100 = 800 observations			. . .
					0 0 0 1 0
		D 8x100 = 800 observations			. . .
					0 0 0 0 1
		E 8x100 = 800 observations			. . .

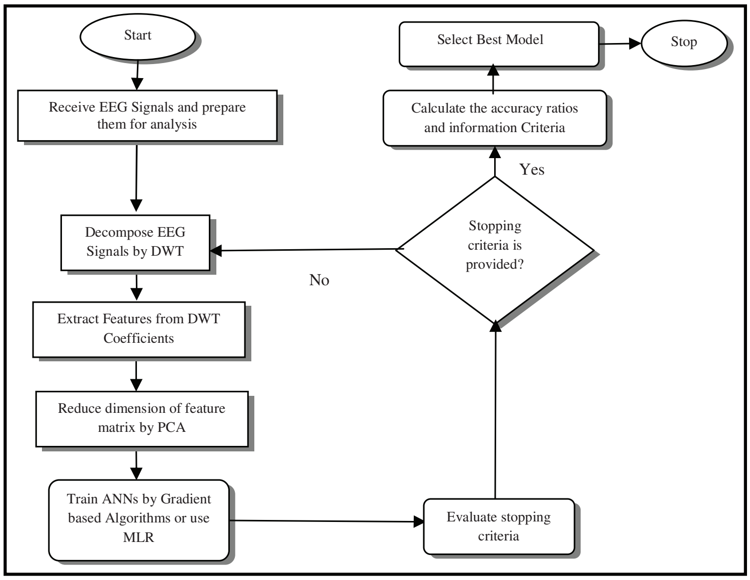

FIGURE 6: The scheme of proposed approach.

TRAINING ANNs USING THE GRADIENT-BASED ALGORITHMS

In the context of epilepsy detection, to improve the classification performance of ANNs, they were trained by various gradient-based algorithms. During the estimation process, the learning algorithms were treated by different tuning parameters whether they increase the performance of ANNs or not. Besides, the over-fitting and complexity of model are handled by early stopping, cross-validation as well as the information criteria. For the cross-validation, the feature matrix was partitioned into three subsets called as training, validation and test. In analysis, three information criteria (AIC, AIC_c and BIC) allow to determine the most efficient number of neurons in the hidden layers. Specifically, the best model configuration is examined by considering the true classification ratios over EEG signals together with MSE and information criteria. According to different training algorithms, all the performances of ANNs are discussed in detail as following.

The Performance of ANNs Trained By The Gradient Descent (GD)

During the model estimation process, to improve the performance of ANNs, GD algorithm was treated by different learning rates. In the hidden layers, the efficient number of neurons was investigated by information criteria, instead of trial and error. The best performances of ANNs are given in Table 3 and Table 4. According to results of Table 3 and Table 4, apparently the information criteria penalizes the complex models due to the excessive number of neurons. As a result, the accurate classification ratios can be obtained over training and test datasets by means of much smaller number of neurons as well. In the Tables, the best model configurations are given with bold font with respect to information criteria and accuracy ratios over test data.

TABLE 3: Performance of GD based DWT.

Neurons	Training data MSE	Test data MSE	AIC	AICc	BIC	Training classes %	Test classes %	Total classes %
6	0.2031	0.2038	-6290.90	-6289.90	-5977.25	99.3	98.5	99.2
7	0.2031	0.2038	-6276.64	-6275.30	-5911.94	99.4	98.8	99.2
12	0.2028	0.2037	-6211.50	-6207.68	-5591.51	99.3	98.8	99.1
15	**0.2025**	**0.2029**	**-6175.64**	**-6169.70**	**-5402.47**	**99.5**	**99.3**	**99.4**
17	0.2025	0.2029	-6147.44	-6139.82	5272.15	99.5	98.8	99.3
20	0.2023	0.2028	-6110.64	-6100.11	-5082.18	99.6	99.3	99.4

TABLE 4: Performance of GD based DWT and PCA.

Features	Variance Explanation %	Neurons	Training data MSE	Test data MSE	AIC	AICc	BIC	Training classes %	Test classes %	Total classes %
20	91.42	9	0.2200	0.2209	-5929.01	-5926.83	-5462.19	91.2	90.5	91.0
30	97.30	15	0.2090	0.2103	-6049.06	-6043.13	-5275.90	96.6	95.3	96.1
33	98.29	15	0.2071	0.2097	-6086.35	-6080.41	-5313.18	97.8	95.5	97.2
39	99.25	13	0.2029	0.2038	-6196.27	-6191.79	-5525.22	99.4	98.8	99.2
42	99.51	10	0.2028	0.2029	-6244.05	-6241.38	-5726.18	99.6	99.3	99.5
49	99.87	8	0.2021	0.2024	-6282.27	-6280.54	-5866.51	99.8	99.8	99.7
57	**99.99**	**7**	**0.2020**	**0.2019**	**-6314.15**	**-6312.80**	**-5949.45**	**100**	**99.8**	**100**
60	100	9	0.2013	0.2018	-6284.55	-6282.37	-5817.73	99.9	99.8	99.9

As seen from Table 3 and Table 4, when the number of neurons in the hidden layer is increased, the performance of training algorithm might improve until a certain number of neurons. However, at the large number of neurons, the model structure gets more complex, so the information criteria penalize it. In addition, the performances of the estimated models decrease slightly. From Table 4, while the number of PCs in the model grows, MSE for training and test data diminish. Also, the satisfaction over information criteria and the classification ratio increase.

Performance of BFGS

BFGS algorithm no needs any initial or tuning parameters during traing of ANNs, because it utilizes the approximations of Hessian matrix. In the hidden layers, the efficient number of neurons was investigated by information criteria, instead of trial and error. The best performances of ANNs are given in Table 5 and Table 6. According to results of Table 5 and Table 6, the information criteria apparently penalize the complex models due to the excessive number of neurons. As a result, the accurate classification ratios can be obtained over training and test datasets by means of much smaller number of neurons as well. In the Tables, the best model configurations are given with bold font with respect to information criteria and accuracy ratios over test data.

As seen from Table 5 and Table 6, when the number of neurons in the hidden layer is increased, the performance of training algorithm might improve until a certain number of neurons. However, at the large number of neurons, the model structure gets more complex, so the information criteria penalize it. In addition, the performances of the estimated models decrease slightly. From Table 6, while the number of PCs in the model grows, MSE for training and test data diminish. Also, the satisfaction over information criteria and the classification ratio increase.

TABLE 5: Performance of BFGS based DWT.

Neurons	Training data MSE	Test data MSE	AIC	AICc	BIC	Training classes %	Test classes %	Total classes %
8	0.2075	0.2092	-6175.85	-6174.11	-5760.09	96.2	95.3	96.1
9	0.2042	0.2067	-6227.37	-6225.19	-5760.55	98.0	96.5	97.6
10	0.2055	0.2088	-6187.87	-6158.20	-5669.99	97.4	95.5	97.0
12	0.2073	0.2065	-6124.44	-6120.62	-5504.45	96.4	97.0	96.2
13	0.2038	0.2041	-6179.25	-6174.77	-5508.20	98.2	98.0	98.0
18	**0.2030**	**0.2074**	**-6123.91**	**-6115.38**	**-5197.57**	**98.6**	**96.3**	**98.3**
20	0.2050	0.2057	-6057.61	-6047.08	-5029.15	97.5	97.0	97.5

TABLE 6: Performance of BFGS based DWT and PCA.

Features	Variance Explanation %	Neurons	Training data MSE	Test data MSE	AIC	AICc	BIC	Training classes %	Test classes %	Total classes %
20	91.42	16	0.2120	0.2163	-5978.75	-5972.00	-5154.52	94.1	91.8	93.5
27	97.45	9	0.2079	0.2096	-6154.09	-6151.91	-5687.28	96.4	95.0	96.0
33	98.29	14	0.2055	0.2061	-6130.31	-6125.13	-5408.20	97.3	96.5	96.8
41	99.44	14	0.2025	0.2038	-6189.34	-6184.16	-5467.23	98.8	98.0	98.6
49	99.87	4	0.2025	0.2025	-6330.26	-6329.79	-6118.73	98.8	98.8	98.8
56	99.99	6	0.2010	0.2011	-6332.20	-6331.19	-6018.55	99.6	99.5	99.5

Performance of Scaled Conjugate Gradient Algorithm (SCG)

SCG algorithm no needs any initial or tuning parameters during traing of ANNs. As known in the literature, it makes very efficient search in the parameter space. In analysis, to improve the performance of ANNs, the efficient number of neurons was investigated by information criteria, instead of trial and error. The best performances of ANNs are given in Table 7 and Table 8. According to results of Table 7 and Table 8, the information criteria apparently penalize the complex models due to the excessive number of neurons. As a result, the accurate classification ratios can be obtained over training and test datasets by means of much smaller number of neurons as well. In the Tables, the best model configurations are given with bold font with respect to information criteria and accuracy ratios over test data.

As seen from Table 7 and Table 8, when the number of neurons in the hidden layer is increased, the performance of training algorithm might improve until a certain number of neurons. However, at the large

TABLE 7: Performance of Scaled Conjugate Gradient-based DWT.

Neurons	Training data MSE	Test data MSE	AIC	AICc	BIC	Training classes %	Test classes %	Total classes %
6	0.2009	0.2010	-6334.36	-6333.36	-6020.71	99.7	99.8	99.6
9	0.2005	0.2009	-6258.38	6254.55	-5638.38	99.7	99.3	99.6
11	0.2002	0.2009	-6277.45	-6274.22	-5708.51	99.9	99.5	99.8
12	0.2005	0.2009	-6258.38	-6254.55	-5638.38	99.8	99.5	99.7
15	0.2003	0.2005	-6219.06	-6213.13	-5445.89	99.8	100	99.8
16	**0.2000**	**0.2000**	**-6211.73**	**-6204.98**	**-5387.51**	**100**	**100**	**100**
20	0.2004	0.2004	-6148.59	-6138.06	5120.13	99.8	99.8	99.7

TABLE 8: Performance of Scaled Conjugate Gradient-based DWT and PCA.

Features	Variance Explanation %	Neurons	Training data MSE	Test data MSE	AIC	AICc	BIC	Training classes %	Test classes %	Total classes %
20	91.42	18	0.2060	0.2105	-6066.11	-6057.58	-5139.76	97.5	94.3	96.4
25	94.90	18	0.2063	0.2082	-6059.14	-6050.61	-5132.79	97.4	96.0	96.5
35	98.73	12	0.2019	0.2047	-6229.13	-6225.30	-5609.13	99.4	97.8	98.9
40	99.37	12	0.2002	0.2013	-6263.32	-6259.49	-5643.32	99.9	99.5	99.8
45	99.69	12	0.2001	0.2009	-6266.09	-6262.27	-5646.10	100	99.8	99.9
55	99.99	12	0.2000	0.2006	-6267.75	-6263.93	-5647.76	100	99.5	99.9
59	~100	12	0.2001	0.2004	-6266.43	-6262.61	-5646.44	100	100	100
63	~100	12	0.2000	0.2004	-6267.06	-6263.23	-5647.06	100	100	100

number of neurons, the model structure gets more complex, so the information criteria penalize it. In addition, the performances of the estimated models decrease slightly. From Table 8, while the number of PCs in the model grows, MSE for training and test data diminish. Also, the satisfaction over information criteria and the classification ratio increase.

Performance of Levenberg-Marquardt Algorithm (LM)

LM algorithm initially needs only a damping parameter μ. Hence, to improve classification performance of ANNs based MSE, LM algorithm were treated with different intial parameter μ's. In analysis, to improve the performance of ANNs, the efficient number of neurons was investigated by information criteria, instead of trial and error. The best performances of ANNs are given in Table 9 and Table 10. According to results of Table 9 and Table 10, the information criteria apparently penalize the complex models due to the excessive number of neurons. As a result, the accurate classification ratios can be obtained over training and test datasets by means of much smaller number of neurons as well. In the Tables, the best model configurations are given with bold font with respect to information criteria and accuracy ratios over test data.

As seen from Table 9 and Table 10, when the number of neurons in the hidden layer is increased, the performance of training algorithm might improve until a certain number of neurons. However, at the large number of neurons, the model structure gets more complex, so the information criteria penalize it. In addition, the performances of the estimated models decrease slightly. From Table 10, while the number of PCs in the model grows, MSE for training and test data diminish. Also, the satisfaction over information criteria and classification ratios increase.

PERFORMANCE OF MULTIARIATE LOGISTIC REGRESSION

In this implementation, classifying EEG signals was handled by MLR based on DWT and PCA. To investigate the significance of estimated models; Cox and Snell, Nagelkerke and McFadden statistics were used as well as Wald test. In the analysis, the feature extraction was made with respect to different window-widths and decomposition levels for DWT. Thus, the model estimation was handled over many data sets

TABLE 9: Levenberg Marquardt based DWT.

Neurons	Training data MSE	Test data MSE	AIC	AICc	BIC	Training classes %	Test classes %	Total classes %
7	0.2007	0.2019	-6323.65	-6322.30	-5958.95	99.7	98.8	99.4
8	0.2003	0.2020	-6317.22	-6315.48	-5901.46	99.9	99.3	99.7
10	**0.2001**	**0.2007**	**-6293.02**	**-6290.35**	**-5775.15**	**100**	**99.8**	**99.9**
13	0.2000	0.2011	-6253.45	-6248.97	-5582.39	100	99.5	99.9
20	0.2002	0.2010	-6151.60	-6141.07	-5123.14	99.9	99.5	99.8

TABLE 10: Performance of Levenberg Marquardt based DWT and PCA.

Features	Variance Explanation %	Neurons	Training data MSE	Test data MSE	AIC	AICc	BIC	Training classes %	Test classes %	Total classes %
20	91.42	14	0.2042	0.2104	-6157.29	-6152.11	-5435.18	98.4	95.3	96.8
27	97.45	15	0.2014	0.2067	-6198.79	-6192.85	-5425.62	99.4	96.8	98.2
35	98.73	15	0.2005	0.2044	-6216.29	-6210.35	-5443.12	99.8	97.8	99.2
42	99.51	6	0.2005	0.2022	-6341.56	-6340.56	-6027.92	99.8	98.8	99.4
45	99.69	16	0.2003	0.2011	-6206.14	-6199.39	-5381.92	99.9	99.5	99.7
55	**99.99**	**8**	**0.2001**	**0.2007**	**-6321.45**	**-6319.71**	**-5905.69**	**100**	**99.8**	**99.9**
60	~100	6	0.2003	0.2011	-6260.79	-6256.97	-5640.80	99.8	99.3	99.7

having different sizes. However, similar to estimation of ANNs, a feature data matrix which was obtained from the automated multi-resolution decomposition using the window-widths with 512 samples at the six level of DWT, was considered as enough to estimate suitable MLR models. Thus, 63 of features were utilized as possible explanatory variables together with log-odds of outcomes. MLR models are estimated over data sets reduced by PCA. The best results were summarized in Table 11.

As seen from Table 11, when the number of PCs is increased in the model, the performance might improve until a certain number of PCs. However, at the high number of PCs, information criteria apparently penalize the complex models even classification ratios increase.

COMPARISON OF THE BEST MODELS TRAINED BY ALL THE CLASSIFIERS

Table 12 displays the best model configurations estimated by ANNs and MLR. From these results, it can be seen that all the models estimated by different approaches produce pretty much accuracy ratios with respect to training data. Nevertheless, their efficiencies seem a little bit decay over test data. Actually, this loss of performance is acceptable, because the test data is not introduced to ANNs before that.

In the literature, as given in Table 13, there exists many remarkable studies in which same epilepsy data set is handled by different approaches. In Table 13, the researchers are interested in different methodologies where they need different frameworks. Hence, Table 13 exhibits only general performances of these methodologies with respect to training or all data sets.

TABLE 11: Performance of MLR based DWT and PCA.

Features	Variance Explanation %	Training data MSE	Test data MSE	AIC	AICc	BIC	Training classes %	Test classes %	Total classes %
20	91.42	0.1650	0.1644	-7167.55	-7167.32	-7021.67	89.7	91.3	89.9
27	95.98	0.1807	0.1793	-6790.10	-6789.69	-6593.16	95.4	94.5	95.3
35	98.73	0.1879	0.1870	-6616.56	-6615.89	-6361.27	98.5	97.8	98.4
40	99.37	0.1953	0.1944	-6453.52	-6452.65	-6161.76	99.9	99.3	99.8
52	**99.95**	**0.1977**	**0.1969**	**-6379.52**	**-6378.07**	**-6000.23**	**100**	**100**	**100**
54	99.98	0.1983	0 1972	-6363.55	-6361.98	-5969.67	100	100	100
59	~100	0.1986	0.1971	-6348.34	-6346.48	-5917.99	100	100	100
61	~100	0.1988	0. 1972	-6339.24	-6337.26	-5894.30	100	100	100
63	100	0.1981	0.1965	-6349.47	-6347.36	-5889.95	100	99.8	100

TABLE 12: Comparison of the best configurations.

Reduction Methods	Algorithms	Features	Neuron Number	AIC	AICc	BIC	Training data MSE	Test data MSE	Training class. %	Test class. %	Total classes %
DWT	GD	63	15	-6175.64	-6169.70	-5402.47	0.2025	0.2029	99.5	99.3	99.4
	BFGS	63	18	-6123.91	-6115.38	-5197.57	0.2030	0.2074	98.6	96.3	98.3
	SCG	63	16	-6211.73	-6204.98	-5387.51	0.2000	0.2000	100	100	100
	LM	63	10	-6293.02	-6290.35	-5775.15	0.2001	0.2007	100	99.8	99.9
	MLR	63	-	-6349.47	-6347.36	-5889.95	0.1981	0.1965	100	99.8	99.9
DWT+PCA	GD	57	7	-6314.15	-6312.80	-5949.45	0.2020	0.2019	100	99.8	100
	BFGS	56	6	-6332.20	-6331.19	-6018.55	0.2010	0.2011	99.6	99.5	99.5
	SCG	59	12	-6266.43	-6262.61	-5646.44	0.2001	0.2004	100	100	100
	LM	55	8	-6321.45	-6319.71	-5905.69	0.2001	0.2007	100	99.8	99.9
	MLR	52	-	-6379.52	-6378.07	-6000.23	0.1977	0.1969	100	100	100

TABLE 13: Existing approaches in literature and their performances.

Author(s)	Method	Dataset	Accuracy (%)
Guler and Ubeyli[10]	Adaptive neuro-fuzzy inference system + DWT	A – E	98.68
Yalcın et al.[11]	DWT+ SVMs	A-E	99.67
Sharma et al.[12]	Time-frequency flexible wavelet transform + SVMs	A-E	100
Bhattacharyya et al.[13]	Q-Wavelet + Entropy + SVMs	A-E	100
Bhati et al.[14]	Cohen–Daubechies–Feauveau biorthogonal filter banks + ANNs	A-E	99.3
Nigam and Graupe[20]	Non-linear preprocessing filter + ANNs	A – E	97.20
Guo et al.[23]	Multiwavelet-entropy features + ANNs	A/B/C/D – E	98.27
Martis et al.[24]	SampEn-DT	A-E	95.7
Amorim et al.[27]	DWT+ ANNs+KNN+RF	A-E	100.00
Ocak[28]	ApEn on DWT coefficients and classifier	A-E	96.00
Wang et al.[29]	Wavelet packet entropy-hierarchical	A-E	99.4
Acharya et al.[30]	CWT+S1+S2+PhEn+Texture-SVM	A-E	96.00
Nicolaou and Georgiou[32]	Permutation entropy and SVMs	A – E	93.55
Kannathal et al.[82]	Entropy + Adaptive neuro-fuzzy inference system	A – E	92.22
Srinivasan et al.[83]	Time-frequency features + Recurrent neural networks	A – E	99.60
Ataee, Avanaki, Shariatpanahi[84]	Wavelet features + ANNs	A – E	94.00
Subasi[85]	Wavelet features + Expert systems	A – E	95.00
Tzallas et al.[86]	Time-frequency analysis + ANNs	A/B/C/D – E	97.73
Polat and Günes[87]	Fourier features + Decision trees	A – E	98.72
Acharya et al.[88]	Non-linear parameters – ApEn, Gaussian	A – E	95.00
Tzallas et al.[89]	Time-frequency analysis and power spectral density + ANNs	A/E – E	100.00
Liang et al[90]	Spectral analysis and principal component analysis + ANNs	A/D – E	98.74
Chua et al.[91]	Magnitude+PhEn, S1, and S2-GMM	A-E	93.1
Orhan et al.[92]	Wavelet features + k-means clustering + ANNs	A – E	99.60
Iscan et al.[93]	Cross correlation and power spectral density + SVMs	A – E	100.00
Yuan et al.[94]	ApEn/Hurst exponent/DFA-SVMs/ANNs	A-E	96.5
Mahajan et al.[95]	PCA, ICA+NN	A-E	93.63
Gandhi et al.[96]	DWT-Spectral entropy+Energy-PNNs	A-E	100.00
Alam and Bhuiyan[97]	Time-frequency analysis and higher order statistics + ANNs	A-E/D – E	100.00
Zainuddin et al.[98]	Wavelet features + ANNs	A – E	98.87
Xie and Krishnan[99]	Wavelet variances + Nearest neighbors	A – E	100.00
Ahammad et al.[100]	DWT-linear classifier	A-E	98.5
Das et al.[101]	Dual-tree complex wavelets + SVMs	A – E	100.00
Chen (2014)[102]	Dual-tree complex wavelet-Fourier features + Nearest neighbors	A – E	100.00
Acharya et al.[103]	ApEn, SampEn, PhEn, S1, and S2 – Fuzzy	A-E	98.1
Xie and Krishnan[104]	DPCA+FFPC+PCPEM	A-E	100
Das et al.[105]	Dual-tree complex wavelets + inverse Gaussian + SVMs	A – E/ AB/CD – E	100.00
Li et al.[106]	DWT based EA+NNE	A-D-E	98.78
Fu et al.[107]	Hilbert marginal spectrum analysis + SVMs	A/B/C/D – E	98.80

DISCUSSION

From the results of analysis, it can be concluded that the automated multi-resolution technique brings out the latent characteristics of epileptic seizures from EEG signals using DWT with db10 mother wavelet. However, the proposed approach can be easily adapted to another mother wavelet in the different cases as well. In the feature extraction process, the features vectors were produced by some important statistics where they are able to explain the behaviors of epileptic seizures over EEG signals. During the estimation procedure, to take into account the complexity of estimated models, RAM and time consumption; the feature matrix was reduced by means of PCA. To improve the classification performance of ANN classifiers, they were trained by different gradient-based algorithms. To handle over/lower fitting during the model estimation process, the feature matrix was separated into three subsets called as training, validation and test. During the estimation process of ANNs, the training algorithms were stopped automatically as soon as the errors over validation data set increases. Essentially, this early detection strategy helps to estimate more robust models. In the hidden layers, the number of neurons is determined by information criteria; thus the excessive complex models are automatically penalized.

In the context of MLR, the multicollinearity is a hypothetical decay and there exists a significant multicollinearity among the feature vectors obtained by DWT at different decomposition levels. To find out this multicollinearity, some robust statistics such as tolerance, VIF, Pearson, Kendall's tau-b and Spearman coefficients were considered. To reduce the dimension of feature matrix, PCA provides remarkable contribution to MLR in terms of complexity and classification performance. To determine the number of components that should enter in the model, the sum of variance explanation rates was used. For instance, according to PCA results, the first 6 principle components account for about 70% of the total variance where these variances are 32.79, 17.35, 7.18, 5.22, 3.99, and 3.82, respectively. However, the models estimated by MLR give better accuracy ratios at the larger variance explanation levels. Therefore, the models were estimated up to 90% variance levels.

From all the analysis results above, it can be said that the performance and robustness of estimated models could be enhanced by feature extraction, DWT and PCA. Also, the number of neurons and PCs play important roles in the classification of EEG signals accurately. Generally, the performance of ANNs depends on many factors such as the network structure, number of neurons, activation functions, learning algorithms and tuning parameters. Conversely, MLR allows more practical procedure to estimate the classification models in terms of controlling and interpreting the system. Unlike ANNs, MLR needs much larger data set to estimate more accurate models according to the results of analysis. Namely, the performance of MLR decreases if their models are estimated over the feature matrices extracted at smaller window-widths and decomposition levels of DWT.

CONCLUSION

Consequently, the proposed procedure is capable of making a comprehensive analysis of EEG signals in the context of epileptic seizures, and estimate reliable and robust models for ANNs and MLR. Especially, these models are able to make more accurate classification of instantaneous EEG signals received from new epileptic signals where they are real seizures or artificially created in the clinical environments. In addition, this approach can be easily applied to another epilepsy data set.

Actually, the proposed approach requires many steps such as the signal decomposition, feature extraction, feature selection, reducing the size of feature matrix, estimation process, selecting the best model configurations, complexity of models, interpreting and discussing the results. Hence, software of devoloped procedure can help users to overcome these challenges exactly. In addition, to estimate more robust and realiable models, more comprehensive data sets are inevitable, but collecting this kind of data is another big challenge for experts due to some legal restrictions. Although there are various benchmark data sets in the literature, they are not enough to attempt an exact time-frequency analysis and figure out all the patterns in the EEG signals. For this reason, this challenge prompts researchers to communicate with neurology institutes in terms of establishing collaboration and making more reliable analysis.

In the future direction, we are planning to focus on developing novel hyrid artificial intelligence approaches. To figure out the performance of these hybrid approaches, they will be applied to more complicated EEG signals with high frequencies and wider time intervals. Thus, this kind of comprehensive data set will allow making deeply time-frequency analysis of EEG signals in the context of epileptic seizures. To do this, we will contact some neurology institutes and clinics, and seeking for a future collaboration. Lastly, we are planning to devolope a user-friendly software by MATLAB GUI.

Source of Finance

This work was supported in part by the Scientific and Technological Research Council of TURKEY (TUBITAK) under grant No. 1059B191401482.

Conflict of Interest

No conflicts of interest between the authors and / or family members of the scientific and medical committee members or members of the potential conflicts of interest, counseling, expertise, working conditions, share holding and similar situations in any firm.

Authorship Contributions

Idea/Concept: *Ezgi Özer, Ozan Kocadağlı;* ***Design:*** *Ezgi Özer, Ozan Kocadağlı;* ***Control/Supervision:*** *Ozan Kocadağlı;* ***Analysis and/or Interpretation:*** *Ezgi Özer, Ozan Kocadağlı;* ***Literature Review:*** *Ezgi Özer, Ozan Kocadağlı;* ***Writing The Article:*** *Ezgi Özer, Ozan Kocadağlı;* ***Critical Review:*** *Ozan Kocadağlı.*

REFERENCES

1. Sharanreddy M, Kulkarni PK. EEG signal classification for epilepsy seizure detection using improved approximate entropy. International Journal of Public Health Science (IJPHS). 2013;2(1):23-32. [Crossref]
2. Zupec-Kania BA, Spellman, E. An overview of the ketogenic diet for pediatric epilepsy. Nutr Clin Pract. 2009;23(6):589-96. PMID: 19033218 [Crossref] [PubMed]
3. Shoeb A, Guttag J. Application of Machine Learning to Epileptic Seizure Detection. Proceedings of the 27th International Conference on Machine Learning. Haifa, Israel. June 21-24, 2010:975-82.
4. Sanei S, Chambers JA. EEG Signal Processing. Chichester, England; Hoboken, NJ: John Wiley & Sons; 2007. p.1-31. [Crossref] [PMC]
5. Taywade SA, Raut RD. A review: EEG signal analysis with different methodologies. Proceedings of the International Journal of Computer Applications. India. 2012:29-31.
6. Agarwal R, Gotman J, Flanagan D, Rosenblatt B. Automatic EEG analysis during long-term monitoring in the ICU. Electroencephalogr Clin Neurophysiol. 1998;107(1):44-58. PMID: 9743272 [Crossref]
7. 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. Phys Rev E Stat Nonlin Soft Matter Phys. 2001;64(6 Pt 1):061907. PMID: 11736210 [Crossref] [PubMed]
8. Mainardi LT, Bianchi LM, Cerutti S. Biosignal Processing: Principles and Practices. In: Liang H, Bronzino JD, Peterson DR, Raton B. NY: CRC Press; 2013. p.7-6.
9. Rosso OA, Figliola A, Creso J, Serrano E. Analysis of wavelet-filtered tonic-clonic electroencephalogram recordings. Med Biol Eng Comput. 2004;42(4):516-23. PMID: 15320461 [Crossref]
10. Güler I, Ubeyli ED. Adaptive neuro-fuzzy inference system for classification of EEG signals using wavelet coefficients. J Neurosci Methods. 2005;148(2):113-21. PMID: 16054702 [Crossref] [PubMed]
11. Yalcın N, Tezel G, Karakuzu C. Epilepsy diagnosis using artificial neural network learned by PSO. Turk J Elec Eng & Comp Sci. 2015;23:421-32. [Crossref]
12. Sharma M, Pachori RB, Acharya UR. A new approach to characterize epileptic seizures using analytic time-frequency flexible wavelet transform and fractal dimension. Pattern Recognition Letters. 2017;9(C):172-9. [Crossref]
13. Bhattacharyya A, Pachori RB, Upadhyay A, Acharya UR. Tunable-Q wavelet transform based multi-scale entropy measure for automated classification of epileptic EEG signals. Appl Sci. 2017;7(4):1-18. [Crossref]
14. Bhati D, Sharma M, Pachori RB, Gadre VM. Time-frequency localized three-band biorthogonal wavelet filter bank using semi-definite relaxation and nonli-

near least squares with epileptic seizure EEG signal classification. Digital Signal Processing. 2017;62(C):259-73. [Crossref]

15. Kumar Y, Dewal ML, Anand RS. Epileptic seizures detection in EEG using DWT-based apen and artificial neural network. Signal, Image and Video Processing (SVIP). 2014;8(7):1323-34. [Crossref]
16. Acharya UR, Sree SV, Suri JS. Automatic detection of epileptic EEG signals using higher order cumulant features. Int J Neural Syst. 2011;21(5):403-14. [Crossref] [PubMed]
17. Acharya UR, Sree SV, Chattopadhyay S, Yu W, Ang PC. Application of recurrence quantification analysis for the automated ıdentification of epileptic EEG signals. Int J Neural Syst. 2011;21(3):199-211. PMID: 21656923 [Crossref] [PubMed]
18. Acharya UR, Oh SL, Hagiwara Y, Tan JH, Adam M, Gertych A, et al. A deep convolutional neural network model to classify heartbeats. Comput Biol Med. 2017;89:389-96. PMID: 28869899 [Crossref] [PubMed]
19. Akin M, Arserim MA, Kiymik MK, Turkoglu I. A new approach for diagnosing epilepsy by using wavelet transform and neural network. Proceedings of the 23rd Annual International Conference of the IEEE Engineering in Medicine and Biology Society. İstanbul, Turkey, October 25-28, 2001:2:1596-9.
20. Nigam VP, Graupe D. A neural-network-based detection of epilepsy. Neurol Res. 2004;26(1):55-60. PMID: 14977058 [Crossref] [PubMed]
21. Güler I, Ubeyli ED. Application of adaptive neuro-fuzzy inference system for detection of electrocardiographic changes in patients with partial epilepsy using feature extraction. Exp Syst Appl. 2004;27(3):323-30. [Crossref]
22. Patnaik LM, Manyam OK. Epileptic EEG detection using neural networks and post-classification. Comput Methods Programs Biomed. 2008;91(2):100-9. PMID: 18406490 [Crossref] [PubMed]
23. Guo L, Rivero D, Dorado J, Rabuñal JR, Pazos A. Automatic epileptic seizure detection in eegs based on line length feature and artificial neural networks. J Neurosci Methods. 2010;191(1):101-9. PMID: 2059503 [Crossref] [PubMed]
24. Martis RJ, Acharya UR, Tan JH, Petznick A, Tong L, Chua CK, et al. Application of ıntrinsic time-scale decomposition (ITD) to EEG signals for automated seizure prediction. Int J Neural Syst. 2013;23(5):1350023. PMID: 23924414 [Crossref] [PubMed]
25. Dehuri S, Jagadev AK, Cho SB. Epileptic seizure identification from electroencephalography signal using de-RBFNS ensemble. Procedia Computer Science. 2013;23:84-95. [Crossref]
26. Rivero D, Aguiar-Pulido V, Fernandez-Blanco E, Gestal M. Using genetic algorithms for automatic recurrent ANN development: an application to EEG signal classification. International Journal of Data Mining, Modeling and Management (IJDMMM). 2013;5(2):182-91. [Crossref]
27. Amorim P, Moraes T, Fazanaroa D, Silvaa J, Pedrini H. Electroencephalogram signal classification based on shearlet and contourlet transforms. Exp Syst Appl. 2017;67(C):140-7. [Crossref]
28. Ocak H. Automatic detection of epileptic seizures in EEG using discrete wavelet transform and approximate entropy. Exp Syst Appl. 2009;36(2):2027-36. [Crossref]
29. Wang D, Miao D, Xie C. Best basis-based wavelet packet entropy feature extraction and hierarchical EEG classification for epileptic detection. Exp Syst Appl. 2011;38(11):14314-20. [Crossref]
30. Acharya UR, Sree SV, Swapna G, Martis RJ, Suri JS. Automated EEG analysis of epilepsy: a review. Knowledge-Based Systems. 2013;45:147-65. [Crossref]
31. Kumari RSS, Jose JP. Seizure detection in EEG using time frequency analysis and SVM. Proceedings of the International Conference on Emerging Trends in Electrical and Computer Technology. Nagercoil, India, March 23-24,2011:626-30.
32. Nicolaou N, Georgiou J. Detection of epileptic electroencephalogram based on permutation entropy and support vector machines. Exp Syst Appl. 2012;39(1):202-9. [Crossref]
33. Bajaj JS, Ridlon JM, Hylemon PB, Thacker LR, Heuman DM, Smith S, et al. Linkage of gut microbiome with cognition in hepatic enceplopathy. Am J Physiol Gastrointest Liver Physiol. 2012;302(1):G168-75. PMID: 21940902 [Crossref] [PubMed] [PMC]
34. Joshi V, Pachori RB. Classification of seizure and nonseizure EEG signals using empirical mode decomposition. IEEE Transactions on Information Technology in Biomedicine. 2012;16(5):1135-42. [Crossref] [PubMed]
35. Joshi V, Pachori RB, Vijesh A. Classification of ictal and seizure-free EEG signals using fractional linear prediction. Biomed Signal Process Control. 2014;9:1-5. [Crossref]
36. Pachori RB, Patidar S. Epileptic seizure classification in EEG signals using second-order difference plot of intrinsic mode functions. Computer Methods Programs Biomed. 2014;113(2):494-502. PMID: 24377902 [Crossref] [PubMed]
37. Fu J, Hou J, Chen L, Wang M, Shen Y, Zhang Z, Bao X. The yeast BDF1 regulates endocytosis via LSP1 under salt stress. Current Microbiology. 2015;70(5):671-8. [Crossref] [PubMed]
38. Sharma R, Pachori RB. Classification of epileptic seizures in EEG signals based on phase space representation of ıntrinsic mode functions. Exp Syst Appl. 2015;42(3):1106-17. [Crossref]
39. Pachori RB. Discrimination between ictal and seizure-free EEG signals using empirical mode decomposition. Research Letters in Signal Processing. 2008:1-5. [Crossref]
40. Tiwari AK, Pachori RB, Kanhangad V, Panigrahi BK. Automated diagnosis of epilepsy using key-point based local binary pattern of EEG signals. IEEE J Biomed Health Inform. 2017;21(4):888-96. PMID: 27416609 [Crossref] [PubMed]
41. Sterman MB, MacDonald LR, Stone RK. Biofeedback training of sensorimotor EEG rhythm in man: effect on epilepsy. Epilepsia. 1974;15(3):395-416. PMID: 4527675 [Crossref] [PubMed]
42. Pfurtscheller G, Flotzinger D, Neuper C. Differentiation between finger, toe and tongue movement in man based on 40 hz EEG. Electroencephalogr Clin Neurophysiol. 1994;90(6):456-60. PMID: 7515789 [Crossref]
43. Ashwal S, Rust R. Child neurology in the 20th century. Pediatr Res. 2003;53(2):345-61. PMID: 12538797 [Crossref] [PubMed]
44. Gao RX, Yan R. Wavelets: Theory and Applications for Manufacturing. New York; London: Springer; 2011. p.49-68. [Crossref]
45. Meyer-Baese A, Schmid V. Pattern Recognition and Signal Analysis in Medical Imaging. 2nd ed. Oxford; 2014. p.71-111. [Crossref]
46. Faust O, Acharya UR, Adeli H, Adeli A. Wavelet-based EEG processing for computer-aided seizure detection and epilepsy diagnosis. Seizure. 2015;26:56-64. PMID: 25799903 [Crossref] [PubMed]

47. Chen D, Wan S, Xiang J, Bao FS. A high-performance seizure detection algorithm based on Discrete Wavelet Transform (DWT) and EEG. PLoS One. 2017;12(3):e0173138. PMID: 28278203 [Crossref] [PubMed] [PMC]
48. Golden RM. Mathematical Methods for Neural Network Analysis and Design. England: The MIT Press; 1996. p. 419.
49. Bishop C. Neural Networks for Pattern Recognition. UK: Oxford University Press; 2010. p.1-27.
50. Kocadagli O, Asıkgil B. Nonlinear time series forecasting with Bayesian neural networks. Exp Syst Appl. 2014;41:6596-610. [Crossref]
51. Akbilgic O, Bozdagan H. A new supervised classification of credit approval data via the hybridized RBF neural network model using information complexity. In: Lausen B, Krolak-Schwerdt S, Böhmer M, eds. Data Science, Learning by Latent Structures, and Knowledge Discovery. Berlin Heidelberg: Springer; 2015. p.13-27. [Crossref]
52. Kocadagli O. A Novel Hybrid Learning Algorithm for Full Bayesian Approach of Artificial Neural Networks. Applied Soft Computing. Elsevier; 2015;35:1-958. [Crossref]
53. Freitas JFG. Bayesian methods for neural networks. PhD. Thesis, UK:Trinity College University of Cambridge and Cambrige University Engineering Department; 2000.
54. Geman S, Bienenstock E, Doursat R. Neural networks and the bias/variance dilemma. Mass Inst Technol. 1992;4(1):1-58. [Crossref]
55. Bozdogan H. Akaike's ınformation criterion and recent developments in ınformation complexity. J Math Psychol. 2000;44(1):62-91. PMID: 10733858 [Crossref] [PubMed]
56. Møller MF. A. scaled conjugate gradient algorithm for fast supervised learning. Neural Networks. 1993;6(4):525-33. [Crossref]
57. Soydaner D, Kocadagli O. Artificial neural networks with gradient learning algorithm for credit scoring. Istanbul University Journal of the School of Business Administration. 2015;44(2):3-12.
58. Sharma S. Applied Multivariate Techniques. America: John Wiley & Sons, Inc; 1996. p.288.
59. Moon TK, Stirling WC. Mathematical Methods and Algorithms for Signal Processing. New Jersey: Prentice Hall; 2000. p.305-20.
60. Armañanzas R, Alonso-Nanclares L, Defelipe-Oroquieta J, Kastanauskaite G, de Sola RG, Defelipe J, et al. Machine learning approach for the outcome prediction of temporal lobe epilepsy surgery. PloS One. 2013;8(4): e62819. PMID: 23646148 [Crossref] [PubMed] [PMC]
61. Jennum P, Pickering L, Christensen J, Ibsen R, Kjellberg L. Morbidity and mortality of childhood - and adolescent - onset epilepsy: a controlled national study. Epilepsy & Behavior. 2017;66:80-5. [Crossref] [PubMed]
62. Jeong A, Wong M. Systemic disease manifestations associated with epilepsy in tuberus sclerosis complex. Epilepsia. 2016;57(9):1443-9. PMID: 27417921 [Crossref] [PubMed]
63. Procházka A, Jech J, Smith, J. Wavelet transform use in signal processing. Proceedings of the 31st International Conference in Acoustic. Prague, Czech Republic. 1994. p.209-13.
64. Daubechies I. Ten Lectures on Wavelets. 2nd ed. Philadelphia: SIAM; 1992. p.109-20. [Crossref] [PubMed]
65. Vernekar K, Kumar H, Gangadharan KV. Computational and experimental approach for fault detection of gears. Journal of Vibration Analysis, Measurement and Control. 2014;2(1):16-29.
66. Kocadagli O, Langari R. Classification of EEG signals for epileptic seizures using hybrid artificial neural networks based wavelet transforms and fuzzy relations. Exp Syst Appl. 2017;88 (C):419-34. [Crossref]
67. Mallat S. A Wavelet Tour of Signal Processing. 2nd ed. San Diego: Academic Press; 1999. p.103-12. [Crossref]
68. Mallat S. A theory for multiresolution signal decomposition: the wavelet representation. IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI). 1989;11(7):674-93. [Crossref]
69. Kociolek M, Materka A, Strzelecki, M, Szczypiński P. Discrete Wavelet Transform- Derived Features for Digital Image Texture Analysis. Proceedings of the International Conference on Signals and Electronic Systems. Lodz, 18-21 September, 2001:163-8.
70. Akansu AN, Serdijn WA, Selesnick IW. Emerging applications of wavelets: a review. Physical Communication. 2010;3(1):11-8. [Crossref]
71. Daubechies I. The wavelet transform, time-frequency localization and signal analysis. IEEE Transactions on Information Theory. 1990;36(5):961-1005. [Crossref]
72. Walker JS. A Primer on Wavelets and Their Scientific Applications. Florida: Chapman & Hall/CRC; 1999. [Crossref]
73. Misiti M, Misiti Y, Oppenheim G, Poggi JM. Wavelet Toolbox 4 User's Guide, 2009. .
74. Hotelling H. Analysis of A Complex of Statistical Variables into Principal Components. Journal of Educational Psychology. 1993:24. [Crossref]
75. Jolliffe IT. Principal Component Analysis. 2nd ed. New York; Springer; 2002. p.112-8.
76. Hastie T, Tibshirani R, Wainwright M. Statistical Learning with Sparsity The Lasso and Generalizations. New York: CRC Press; 2016. p.202-12. [Crossref]
77. Mehrjoo S, Bashiri M. An application of principal component analysis and logistic regression to facilitate production scheduling decision support system: an automotive ındustry case. Journal of Industrial Engineering International. 2013;9:1-12. [Crossref]
78. Zaki MJ, Meira W Jr. Data Mining and Analysis: Fundamental Concepts and Algorithms. New York: Cambridge University Press; 2014. p.208-21. [Crossref]
79. Hladnik A. Image compression and face recognition: two ımage processing applications of principal component analysis. International Circular of Graphic Education and Research. 2013;6:56-61.
80. Hosmer DW, Lemeshow S. Applied Logistic Regression. 2nd ed.New York: John Wiley & Sons Inc; 2000. p.7-11. [Crossref] [PubMed]
81. James G, Witten D, Hastie T, Tibshirani R. An Introduction to Statistical Learning with Applications in R. New York: Springer; 2013. p.93-8. [Crossref]
82. Kannathal N, Choo ML, Acharya UR, Sadasivan PK. Entropies for detection of epilepsy in EEG. Comput Methods Programs Biomed. 2005;80(3):187-94. PMID: 16219385 [Crossref] [PubMed]
83. Srinivasan V, Eswaran C, Sriraam N. Artificial neural network based epileptic detection using time-domain and frequency-domain features. J Med Syst. 2005;29(6):647-60. PMID: 16235818 [Crossref] [PubMed]
84. Ataee P, Avanaki AN, Shariatpanahi HF, Khoee SM. "Ranking Features of Wavelet-Decomposed EEG Based on Significance in Epileptic Seizure Prediction," Proceedings of the 14th European Signal Processing Conference. Florence, Italy, September 4-8, 2006. p.1-4.
85. Subasi A. EEG signal classification using wavelet feature extraction and a mixture of expert model. Exp Syst Appl. 2007;32(4):1084-93. [Crossref]

86. Tzallas AT, Tsipouras MG, Fotiadis DI. Automatic Seizure Detection based on Time-Frequency Analysis and Artificial Neural Networks. Comput Intell Neurosci. 2007:80510. PMID: 18301712 [Crossref] [PubMed] [PMC]
87. Polat K, Gunes S. Classification of epileptiform EEG using a hybrid system based on decision tree classifier and fast Fourier transform. Applied Mathematics and Computation (JAMC). 2007;187(2):1017-26. [Crossref]
88. Acharya UR, Chua KC, Lim TC, Dorithy, Suri JS. Automatic identification of epileptic EEG signals using nonlinear parameters. J Mech Med Biol. 2009;9(4):539-53. [Crossref]
89. Tzallas AT, Tsipouras MG, Fotiadis DI. Epileptic seizure detection in EEGs using time-frequency analysis. IEEE Trans Inf Technol Biomed. 2009;13(5):703-10. PMID: 19304486 [Crossref] [PubMed]
90. Liang SF, Wang HC, Chang WL, Combination of EEG complexity and spectral analysis for epilepsy diagnosis and seizure detection. EURASIP J Adv Signal Process. 2010;2010:853434. [Crossref]
91. Chua KC, Chandran V, Acharya UR, Lim CM. Application of higher order spectra to identify epileptic EEG. J Med Syst. 2011;35(6):1563-71. PMID: 20703761 [Crossref] [PubMed]
92. Orhan U, Hekim M, Ozer M. EEG Signals classification using the K-means clustering and a multilayer perceptron neural network model. Exp Syst Appl. 2011;38:13475-81. [Crossref]
93. Iscan Z, Dokur Z, Demiralp T. Classification of electroencephalogram signals with combined time and frequency features. Exp Syst Appl. 2011;38(8):10499-505. [Crossref]
94. Yuan Q, Zhou W, Li S, Cai D. Epileptic EEG classification based on extreme learning machine and nonlinear features. Epilepsy Res. 2011;96(1-2):29-38. PMID: 21616643 [Crossref] [PubMed]
95. Mahajan K, Vargantwar MR, Rajput SM. Classification of EEG using PCA, ICA and neural network. International Journal of Engineering and Advanced Technology (IJEAT). 2011;1(1):80-3.
96. Gandhi T, Panigrahi BK, Bhatia M, Anand S. Discrete harmony search based expert model for epileptic seizure detection in electroencephalography. Exp Syst Appl. 2012;39(4):4055-62. [Crossref]
97. Alam SM, Bhuiyan MI. Detection of seizure and epilepsy using higher order statistics in the EMD domain. IEEE J Biomed Health Inform. 2013;17(2):312-8. PMID: 24235109 [Crossref] [PubMed]
98. Zainuddin Z, Huong LK, Pauline O. Reliable epileptic seizure detection using an ımproved wavelet neural network. Australas Med J. 2013;6(5):308-14. PMID: 23745153 [Crossref] [PubMed] [PMC]
99. Xie S, Krishnan S. Wavelet-based sparse functional linear model with applications to EEGs seizure detection and epilepsy diagnosis. Med Biol Eng Comput. 2013;51(1-2):49-60. PMID: 23054383 [Crossref] [PubMed]
100. Ahammad N, Fathima T, Jospeh P. Detection of epileptic seizure event and onset using EEG. Biomed Res Int. 2014; 2014:450573. PMID: 24616892 [Crossref] [PubMed] [PMC]
101. Das AB, Bhuiyan MIH, Alam SS. A Statistical Method for Automatic Detection of Seizure and Epilepsy in The Dual Tree Complex Wavelet Transform Domain. Proceedings of the IEEE In International Conference on Informatics, Electronics and Vision. Dhaka, Bangladesh, May 23-24, 2014:1-6. [Crossref]
102. Chen G. Automatic EEG seizure detection using dual-tree complex wavelet-fourier features. Exp Syst Appl. 2014;41(5):2391-4. [Crossref]
103. Acharya UR, Molinari F, Sree SV, Chattopadhyay S, Ng KH, Suri JS. Automated diagnosis of epileptic EEG using entropies. Biomed Signal Process Control. 2012;7(4):401-8. [Crossref]
104. Xie S, Krishnan S. Dynamic principal component analysis with non-overlapping moving window and its applications to epileptic EEG classification. Sci World J. 2014;2014:419308. [Crossref] [PubMed] [PMC]
105. Das AB, Bhuiyan MIH, Alam SS. Classification of EEG signals using normal ınverse gaussian parameters in the dual-tree complex wavelet transform domain for seizure detection. Signal, Image and Video Processing (SVIP). 2016;10(2):259-66. [Crossref]
106. Li M, Chen W, Zhang T. Classification of epilepsy EEG signals using DWT-based envelope analysis and neural network ensemble. Biomed Signal Process Control. 2017;31:357-65. [Crossref]
107. Fu K, Qu J, Chai Y, Zou T. Hilbert marginal spectrum analysis for automatic seizure detection in EEG signals. Biomed Signal Process Control. 2015;18:179-85. [Crossref]