Cognitive Neurodynamics (2018) 12:271–294
https://doi.org/10.1007/s11571-018-9477-1

RESEARCH ARTICLE

Detecting epileptic seizure with different feature extracting strategies using robust machine learning classification techniques by applying advance parameter optimization approach

Lal Hussain[1,2]

Received: 9 October 2017 / Revised: 1 December 2017 / Accepted: 18 January 2018 / Published online: 25 January 2018

Abstract

Epilepsy is a neurological disorder produced due to abnormal excitability of neurons in the brain. The research reveals that brain activity is monitored through electroencephalogram (EEG) of patients suffered from seizure to detect the epileptic seizure. The performance of EEG detection based epilepsy require feature extracting strategies. In this research, we have extracted varying features extracting strategies based on time and frequency domain characteristics, nonlinear, wavelet based entropy and few statistical features. A deeper study was undertaken using novel machine learning classifiers by considering multiple factors. The support vector machine kernels are evaluated based on multiclass kernel and box constraint level. Likewise, for K-nearest neighbors (KNN), we computed the different distance metrics, Neighbor weights and Neighbors. Similarly, the decision trees we tuned the paramours based on maximum splits and split criteria and ensemble classifiers are evaluated based on different ensemble methods and learning rate. For training/testing tenfold Cross validation was employed and performance was evaluated in form of TPR, NPR, PPV, accuracy and AUC. In this research, a deeper analysis approach was performed using diverse features extracting strategies using robust machine learning classifiers with more advanced optimal options. Support Vector Machine linear kernel and KNN with City block distance metric give the overall highest accuracy of 99.5% which was higher than using the default parameters for these classifiers. Moreover, highest separation (AUC = 0.9991, 0.9990) were obtained at different kernel scales using SVM. Additionally, the K-nearest neighbors with inverse squared distance weight give higher performance at different Neighbors. Moreover, to distinguish the postictal heart rate oscillations from epileptic ictal subjects, and highest performance of 100% was obtained using different machine learning classifiers.

Keywords Support vector machine · Decision tree · Ensemble classifier · K-nearest neighbors · Classification · Epilepsy · Seizure detection

Introduction

Epilepsy is one of the most common neurological disorder with a prevalence of 1–2% of the world's population (Mormann et al. 2007). There is prevalence of sudden unexpected death in epilepsy (SUDEP) due to unexpected, unwitnessed, nondrowning, nontraumatic death in patients suffered from epilepsy with or without evidence of seizure. About 50 million of people in the world are severely affected by the neurological disorder called epilepsy. It is the second common neurological disorder after stroke (Subasi et al. 2017).

Epileptic seizure occurs due to the sudden malfunctioning and synchronization of set of neurons thereby reflecting the excessive and hyper synchronous activity of neurons in the brain. The recurrent seizures also known as epileptic seizures are the hallmark of epilepsy. According to the clinical manifestation, these seizures are divided into generalized, focal, partial, unclassified and unilateral

✉ Lal Hussain
Lall_hussain2008@live.com; Lal.hussain@ajku.edu.pk

[1] Quality Enhancement Cell (QEC), The University of Azad Jammu and Kashmir, City Campus, Muzaffarabad, Azad Kashmir 13100, Pakistan

[2] Department of Computer Science and IT, The University of Azad Jammu and Kashmir, City Campus, Muzaffarabad 13100, Pakistan

(Tzallas et al. 2009; James and Eng 1997). During the focal epileptic seizures, only part of hemisphere is affected and these seizures produce symptoms in corresponding body parts or in some related mental functions. Moreover, during the generalized epileptic seizures, the entire body is involved and bilateral motor symptoms are produced with the loss of consciousness. These types of seizure can occur at any age stage and are more prominent in younger and older demographics (Hassan et al. 2016). How to predict and diagnose the epileptic seizures more effectively is still a challenging task for the researchers.

For detecting epileptic seizures and spikes, some traditional methods such as visual scanning of EEG recordings have been used which are expensive, inaccurate and time consuming taking several days to complete (Ocak 2009). To cope up with this problem, it is required to develop most robust and promising techniques to detect the epileptic activity in EEG signals (Hassan et al. 2016; Tzallas et al. 2007; Guo et al. 2011; Fu et al. 2015). Recently, researchers used several automated methods to detect the epileptic activity in the brain (Ocak 2009; Fu et al. 2015). Moreover, researchers (Gotman 1982; Gotman and Gloor 1976) employed automated detection methods to detect the EEG epileptic seizures. Adeli et al. (2003) employed wavelet transform to study the dynamics in EEG epileptic signals. Ican et al. (2011) extracted time and frequency features to classify the EEG epileptic signals. Guo et al. (2010) used multiwavelet transform to detect the epileptic signals and extracted features using approximate entropy and artificial neural networks for classification purposes. Moreover, (Tzallas et al. 2007) employed time–frequency, (Subasi 2007) wavelet, (Kannathal et al. 2005) entropy for features extraction and artificial neural networks (Tzallas et al. 2007; Srinivasan et al. 2007; Nigam and Graupe 2004) for classification and detection of epileptic seizures. Orhan et al. (2011) applied K-mean clustering and artificial neural networks for classification of EEG epileptic signals. Likewise, spectral analysis was performed by Kang et al. (2015), mixed band wavelet by Ghosh-Dastidar et al. (2007) and wavelet transform and artificial neural networks are proposed by Ocak (2009) for automatic detection of epileptic seizures in EEG signals.

Most recently, researchers used different classification methods such as Bashar et al. (2016) employed multivariate EMD and short time Fourier transform and Bashar et al. (2015) used dual tree complex wavelet transform to classify the EEG left and right-hand movement. Moreover, automatic sleep detection and classification was made by extracting statistical and spectral features, Hassan (2015, 2016) and Stochholm et al. (2016) used adaptive boosting classifier, Hassan (2016) extracted statistical features from infinite sum of intrinsic mode function (IMF) using different machine learning classifier. Likewise, Hassan and Bhuiyan (2016) extracted spectral features using complex tree wavelet transform, Hassan and Haque (2016) statistical and spectral features using Bootstrap aggregating, Hassan and Bhuiyan (2016, 2017) extracted statistical features from EMD using ensemble methods and extreme machine learning. Similarly, Hassan and Bhuiyan (2017) employed normal inverse Gaussian parameter and adaptive boosting, Hassan and Subasi (2017) tunable Q wavelet transform [TQWT], Hassan and Bhuiyan (2017), Hassan and Haque (2017) used ensemble empirical mode decomposition and random under sampling boosting (Fig. 1).

In this paper, we employed automated techniques to detect the epileptic activity in EEG signals by extracting different feature extracting strategies such as multi domain features, nonlinear dynamical measures with fast entropy measure using KD tree algorithmic approach and wavelet entropy in order to improve the detection accuracy. Moreover, we have employed most robust machine learning techniques by tuning the parameters that outer performed than the existing techniques. The time domain features are extracted which are most widely used in variety of applications such as heart and brain for variability analysis. The frequency domain features are extracted used in many applications for spectral analysis to capture the frequency component. The features from nonlinear dynamics are extracted such as approximate entropy and sample entropy with improved performance using KD tree algorithm approach, which is more efficient with respect speed and memory performance. The wavelet entropy features comprised of Shannon entropy, threshold, log energy and sure. Likewise, additional statistical features are also extracted such as smoothness, skewness, kurtosis and root mean square with mean. The combination of features is used as input the novel machine learning classifiers. For support vector classification kernel choice, we employed multiscale kernels and box constraint levels that improves the evaluation performance from default parameters. For K-nearest neighbors, we applied different distance calculation metrics based on Euclidean distance, city block, Chebyshev, cubic, cosine, correlation, spearman, hamming and jaccard. Likewise, for KNN, the performance was also evaluated based on different Neighbors and distance weight i.e. equal, inverse and squared inverse. The performance evaluation varied by changing these set of parameters. For Decision trees, the basic parameters are tuned based on maximum number of splits and split criteria etc. Likewise, performance was also evaluated using different ensemble methods with number of learning and learning rate.

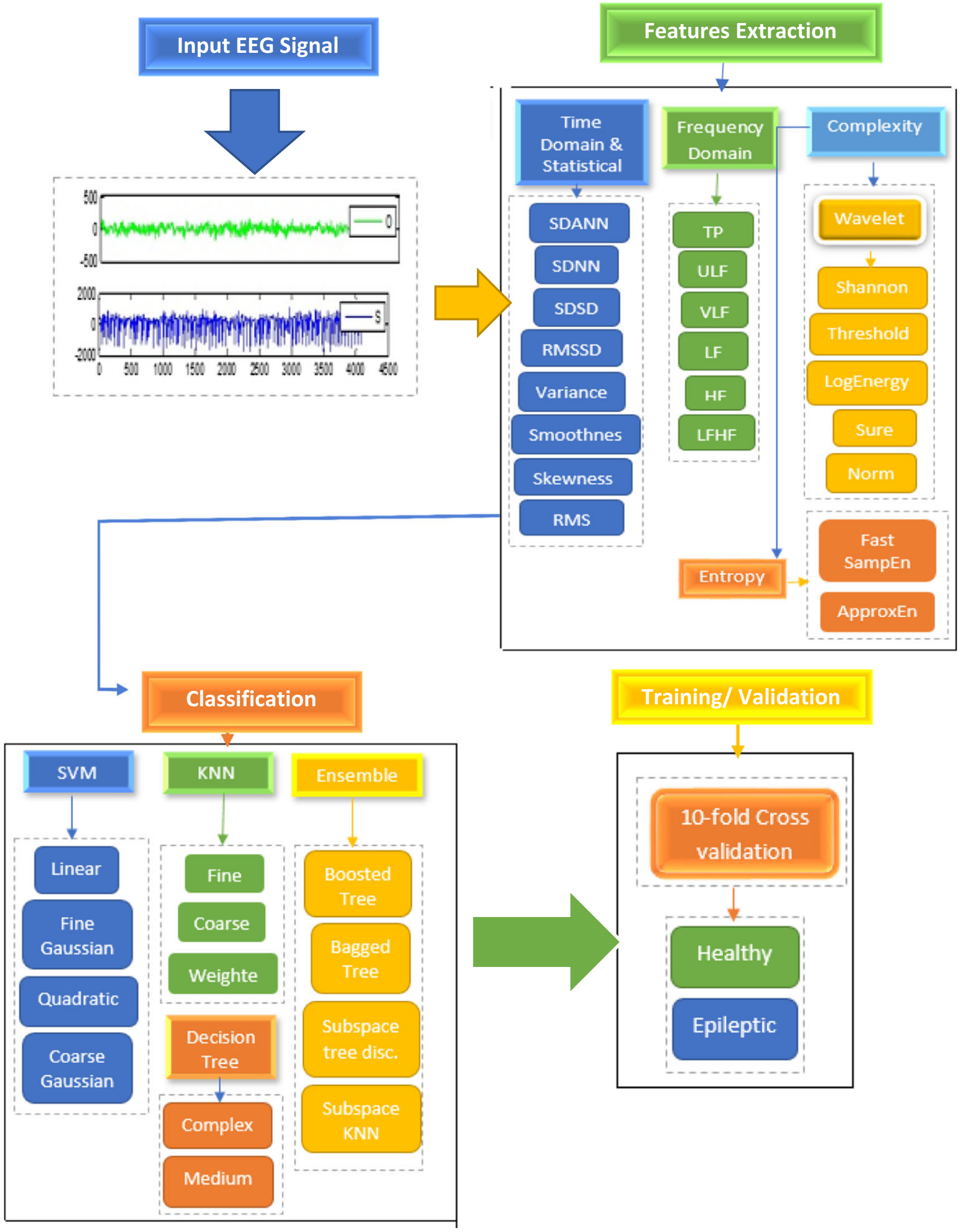

Fig. 1 Schematic diagram of epileptic seizure detection by extracting different feature extracting strategies and multi-criterion machine learning classifiers

Materials and methods

The data set was taken from a publicly available database (http://www.meb.uni-bonn.de/epileptologie/science/physik/eegdataold.html), made available by department of epileptology, Bonn University and its detailed description is provided by Andrzejak et al. (2001). The EEG data was recorded with 128-channel amplifier system, using an average common reference. The data were digitized at sampling frequency of 173.61 Hz using 12-bit resolution. The spectral bandwidth of the data acquisition system varied from 0.5 to 85 Hz. The whole EEG data Comprised of five sets (denoted as Z, O, N, F and S), each containing 23.6 s duration 100 single-channel EEG segments. The sets O and Z were recorded from five healthy volunteers during awake state with eye closed (set O) and eye open (set Z) using a standardized electrode placement scheme. Sets N, F and S originated from an EEG archive of pre-surgical diagnosis. Segments in set F were acquired from the epileptogenic zone and those in set N were acquired from the hippocampal formation of the opposite hemisphere of the brain. The segments in sets N and F contained EEG recordings acquired during seizure free intervals and segments in set S contained seizure activity. For deeper study, we used EEG signals from healthy subjects Set O and epileptic subject Set S.

Another data of Postictal heart rate oscillations with partial epilepsy dataset were taken from publicly available database (Goldberger et al. 2000; Al-Aweel et al. 1999) which consists of 11 partial seizures recorded in 5 female patients during continuous EEG/ECG/video monitoring. The ages of patients ranged from 31 to 48 years with no clinical evidences of cardiac disease and have partial seizures with or without evidence of secondary generalization from frontal or temporal foci. An approved protocol by Israel Deaconess Medical Center was used to made the recordings customized software were used to analyse the data offline. The experienced electroencephalographer (D.L.S) masked with respect to heart rate variability analysis visually identified onset and offset seizures to the nearest 0.1 s. The continuous signal ECG signals were sampled at 200 Hz. The heart beat annotations were obtained using the commercially available arrhythmia analysis software (Ho et al. 1997). The higher order nonstationary fluctuations in heart rate that could mask low frequency oscillations were removed. The time series was detrended using least square filter fourth degree polynomial. The spectral density was estimated using fast Fourier transform technique with rectangular window.

Biological signals represent the pattern of change of a living system that are critical to understand the dynamics of healthy biological systems and how pathological disturbance or aging affects their robustness. With the advent of modern computerized data acquisition system, advanced monitoring devices and instrumentational technologies, the field of biomedical signal processing is growing its popularity. Considerable interest among biomedical researchers has been found to develop innovative methods and tools for biomedical signal processing during last two decades.

The clinicians have recognized the physiological rhythmic alterations which are associated with the disease. The physicians make use of human eye as an excellent pattern recognition device which is capable of interpreting of highly complex ECG and EEG signals (Glass and Kaplan 1993). However, sophisticated analysis of variability provides a quantity of integrity of underlying systems that produces the dynamics. The nature of the system can be best described by the spatial and temporal organization of a complex system i.e. change in patterns of interconnection in term of connectivity and variation in pattern over time in term of variability—represent very important means with which to treat and prognosticate the patients (Glass and Kaplan 1993; Seely and Macklem 2004). Based on the variable and nonlinear and nonstationary dynamics, the multimodal features have been extracted from the healthy, epileptic seizure with ictal intervals and postictal heart rate oscillations.

To measure the variability, various factors are altered e.g. head up tilt or standing (sympathetic activity is increased), deep breathing (respiratory rate induced HRV is increased) that alter the HRV indices in healthy individuals (Seely and Macklem 2004). Moreover, the brain signals variability (i.e. temporal transit fluctuations in brain signals) conveys important information about network dynamics. In complex dynamical systems, such as brain, the variability facilitates the transitions between possible functional network configurations in absence or presence of external stimulation (Lewis and Bates 2013).

Likewise, the physiological signals are considered as sum of sinusoidal oscillations with distinct frequencies. In this regard, conversion is made from time domain to frequency domain using Fourier transform (Takeda et al. 1982) developed by a mathematician in 1807. For conversion from time domain to frequency domain, Wavelet and Hilbert transforms (Gabor 1946) are most commonly used methods as well. In this way, the amplitude of each cosine and sine wave can be determined as function of its frequency also known as spectral analysis because it provides an evaluation of power (amplitude) of contributing frequencies to the underlying signal.

Time domain analysis technique have been used for statistical analysis in term of standard deviation (SDNN) to measure the global variation, standard deviation of average interval (SDANN) to measure the long-term variations, square root of mean squared differences of consecutive NN intervals (RMSSD) to evaluate short term variations. These time domain measures have extensively been used in heart rate variation and brain variations such as coronary artery disease (Rich et al. 1988; Van Hoogenhuyze et al. 1991) to diagnose the increased mortality risk in patients, congestive heart failure (Bilchick et al. 2002; Ponikowski et al. 1997), dilated cardiomyopathy (Tuininga et al. 1994) and post infarction patients (Bigger et al. 2016; Casolo et al. 1992; Kleiger et al. 1987). Moreover, for detecting epileptic seizure several automated methods in time domains (Ocak 2009; Fu et al. 2015; Lee et al. 2014), Fourier spectral analysis for extracting features (Polat and Güneş 2007), frequency domain methods (Polat and Güneş 2007), DWT-based methods (Faust et al. 2015), and fast-Fourier transforms (Tzallas et al. 2007, 2009) were employed.

Signal pre-processing

The low and weak frequency signals are usually suffered from complex low frequency noise such as system interference. Before detecting and analysis of epileptic seizure, EEG signals preprocessing and noise removal is applied (Rajendra Acharya et al. 2012; Acharya et al. 2015). In this work, the wavelet threshold denoising method is used that provide better performance than Fourier transform denoising method (Gajic et al. 2015). The fourth order Daudechies (db4) wavelet was selected that is more suitable for nonstationary signals (Stanley Raj et al. 2016) due to its good local approximation performance. The Principal component analysis was also applied for dimensionality reduction to remove the irrelevant features. For denoising the nonstationary signals, following wavelet threshold method (Dragotti and Vetterli 2003) was applied:

$$\lambda = \delta\sqrt{2logN}$$

where λ is the wavelet threshold, δ is the standard deviation and N is the length of the sample signal respectively. The wavelet threshold method is most robust to remove the white noise buried in the EEG signal (Walters-Williams and Li 2011).

Features extraction

Features extraction is one of the most important step before applying the Machine Learning and Neural networks classification techniques for detection and prediction purposes. It requires an optimum feature set that should effectively discriminate the subjects. Features extraction is solely specific to the problem. Rathore et al. (2014) and Ferland et al. (2017) extracted hybrid and geometric features for automatic colon detection of cancer. Dheeba et al. (2014) extracted texture features for breast cancer detection. Hussain et al. (2014) computed texture and morphological features to detect and classify the human face from non-faces. Moreover, Hussain et al. (2017a) recently extracted acoustic features such as volume, pitch, prosodic features as frequency minimum, maximum, sum, Mel frequency cepstral coefficients to emotion recognition in human speech. They also extracted time and frequency based features to detecting heart rate and heart rate variability.

In this study, multimodal feature extracting strategy was used for detecting the epileptic seizure. The EEG signals are highly complex having nonlinear and nonstationarity behavior. Qu et al. applied empirical mode decomposition (EMD) to decompose EEG signals into a collection of intrinsic mode functions (IMFs), which has the ability to quantify the dynamics of nonstationary and nonlinear processes (Fu et al. 2015). IMF features extracted from EMD are used as input to the Support Vector Machine to detect and classify the epileptic seizure. The results obtained from IMFs gives high classification accuracy. However, few studies (Wang et al. 2017) suggest the use of multimodal features i.e. combing features from multi-domain alongwith the nonlinear features for classifying the epileptic seizure. This will give a unified framework to include the advantages of varying characteristics of EEG signals. In this study, we also have extracted the features based on time domain, frequency domain, complexity based measures and wavelet entropy methods for classifying the epileptic seizure subjects from that of healthy subjects and postictal heart rate oscillations. Apart, in this work, we extracted nonlinear features using sample entropy based on KD tree algorithmic approach (fast Sample entropy) and approximate entropy which gives outer performance than results obtained by Wang et al. (2017) and are consistent with the results obtained by Hussain et al. (2017b). Recently, Hussain et al. (2017b) and Pan et al. (2011) employed fast MSE which gives statistically more effective results than traditional MSE with reduced computational and memory complexity.

Linear methods

To measure the variability in physiological signals (i.e. ECG or EEG etc.) affected by different pathologies, the time and frequency domain techniques are most widely used to capture the time variations and spectral information. Detailed analysis techniques developed to characterize the variability analysis of EEG and ECG signals have been described in detail by Seely and Macklem (2004), Malik (1996) for patients suffered from different variability dysfunctions (Esco et al. 2017; Choi and Shin 2017; Geronikolou et al. 2017; Sima et al. 2017; Kuang et al. 2017; Fujita et al. 2016; Dodds et al. 2017) including heart rate variability, breathing, depression, pulse variability, insomnia problems and epilepsy etc. Ayoubian et al. (2013) extracted different time, frequency and time–frequency domain features to detect the epileptic seizure. They extracted the HF activities and peak, wavelet entropy that contains the important characteristics of seizure onset patterns in HF activities. Moreover, the low frequency ranges (1–70 Hz) also contains the most relevant information pertinent to detect the epileptic seizure. The following time and frequency domain features are extracted from healthy and epileptic patients and postictal heart rate oscillations in patients suffered from epilepsy.

1. Time domain analysis

Different parameters can be extracted by time domain analysis of the segment and time domain methods are the simplest to perform

SDSD: standard deviation of differences between adjacent time series intervals in each segment.
SDNN: standard deviation of the consecutive intervals in each segment.

$$\mathrm{SDNN} = \sqrt{\frac{1}{N-1}\sum_{j=1}^{N}\left(EEG_j - \overline{EEG}\right)^2} \tag{1}$$

RMSSD: Is the square root of the mean squared differences of N successive EEG time series intervals

$$\mathrm{RMSSD} = \sqrt{\frac{1}{N-1}\sum_{j=1}^{N-1}\left(EEG_{j+1} - \overline{EEG_j}\right)^2} \tag{2}$$

SDANN: Standard deviation of the averages of EEG intervals.

$$\mathrm{SDANN} = \mathrm{SD}[u_1, u_2, u_3, u_4 \ldots u_n] \tag{3}$$

2. Frequency domain analysis

Time domain methods are simple but they do not have the ability to discriminate between sympathetic and parasympathetic contributions of HRV and EEG variations.

Total power (TP): Total spectral power of all EEG time series intervals up to 0.4 Hz.
Very low frequency (VLF): Total spectral power of all EEG consecutive intervals between 0.003 and 0.04 Hz.
Low frequency (LF): Total spectral power of all EEG consecutive between 0.04 and 0.15 Hz.
High frequency (HF): Total spectral power of all EEG consecutive intervals between 0.15 and 0.4 Hz.
ULF: Total spectral power of all EEG consecutive intervals up to 0.003 Hz.
LF/HF Ratio: Ratio of low to high frequency power. This measures overall balance between sympathetic and parasympathetic systems.

Nonlinear methods

Biological signals are basically the output of multiple interacting components of a biological system exhibiting complicated patterns. These patterns of change may contain useful hidden information about the dynamics of these systems. It is unrealistic to extract valuable information using traditional data analysis techniques. Following are most commonly used complexity base measure.

Approximate entropy

Approximate entropy (ApEn) developed by Pincus (1991) is a statistical measure used to quantify the regularities in data. It shows the probability that similar observation patterns do not repeat.

$$\mathrm{ApEn}(\mathrm{m}, \mathrm{r}\cdot \mathrm{N}) = \emptyset^m(r) - \emptyset^{m+1}(r) \tag{4}$$

Fast sample entropy with KD tree approach

(SampEn) Sample entropy suggested by Costa et al. (2002) is a modified form of approximate entropy. It is used to assess the physiological time series signal. Sample entropy when comparing with approximate entropy shows good features like independent data length and trouble-free implementation. It can easily be implemented in many programming languages.

Thus, sample entropy can be more precisely computed using following formula:

$$\mathrm{SampEn}(\mathrm{m}, \mathrm{r}) = \lim_{\mathrm{N}\to\infty} -\ln\frac{P^m(r)}{Q^m(r)} \tag{5}$$

where $P^m(r)$ denotes the probability that two sequences will still match for m + 1 points and $Q^m(r)$ is the probability that two sequences will matches for m points (with tolerance of τ); where self matches are excluded. In this regard Eq. (1) can be expressed as:

$$\mathrm{SampEn}(\mathrm{m}, \mathrm{r}, \mathrm{N}) = -\ln\frac{P^m(r)}{Q^m(r)} \tag{6}$$

By setting $Q = \left\{\frac{[(N-m-1)(N-m)]}{2}\right\}Q^m(r)$ and $P = \left\{\frac{[(N-m-1)(N-m)]}{2}\right\}P^m(r)$

We have $\frac{P}{Q} = \frac{P^m(r)}{Q^m(r)}$ and thus sample entropy can be expressed as:

$$\mathrm{SampEn}(\mathrm{m}, \mathrm{r}, \mathrm{N}) = \frac{P^m(r)}{Q^m(r)} \tag{7}$$

where P is the total number of forward matches of length m + 1 and Q is the total number of template matches of length m. Here we used sample entropy with KD tree algorithmic base approached as implemented by Hussain et al. (2017b) which provide improved performance and is more effective with respective to time and space complexity.

Wavelet entropy

In signal processing applications entropy is commonly used for analysis of nonlinear time series data. Commonly used entropy methods (Wang et al. 2011) include Shannon, Log Energy, Threshold, Sure and Norm etc. Shannon entropy (Wang et al. 2011) was employed to measure the complexity of signal to wavelet coefficients generated by WPT where larger values show high uncertainty process and therefore higher complexity. Wavelet entropy used by Rosso et al. (2001) which provided the useful information to measure the underlying dynamical process associated with the signal. The entropy 'E' must be an additive information cost function such that E $(0) = 0$ and $\mathrm{E}(\mathrm{S}) = \sum_i \mathrm{E}(\mathrm{S_i})$.

Shannon entropy

Shannon entropy (Wu et al. 2013) was first proposed in 1948 named after Claude Shannon. Since then, it is most extensively applied in the information sciences. Shannon entropy is a measure of the uncertainty associated with a random variable. Specifically, Shannon entropy quantifies the expected value of the information contained in a message. The Shannon entropy of a random variable X can be defined as follow:

$$V(X) = V(P_1, \ldots, P_n,) = -\sum_{i=1}^{n} P_i log_2 P_i \quad (8)$$

$$P_i = P_r(X = x_i) \quad (9)$$

where Pi is defined in Eq. (9) with x_i indicating the ith possible value of X out of n symbols, and Pi denoting the possibility of $X = x_i$.

Wavelet norm entropy

Wavelet Norm entropy (Avci et al. 2007) is defined as:

$$\mathrm{E(S)} = \frac{\sum_i |\mathrm{S_i}|^p}{\mathrm{N}} \quad (10)$$

where p is the power and must be $1 \ll \mathrm{P} < 2$ the terminal node signal and (si) i the waveform of terminal

Classification

The robust machine learning algorithms were employed to detect and predict the epileptic seizure. The accuracy, and other performance evaluation parameters are estimated using this model. The known label of the test sample is compared with the classified results from the model. Ten-fold cross validation was used for training, test and validation purposes.

In machine learning, the support vector machine (SVM) classifier is most well-known supervised learning method using finite sample theory (Burges 1998). Based on empirical error minimization, the traditional methods in small sample cases are prone to generate the overfitting problem, while SVM is based on structural risk minimization principle and has the good generalization ability (Li et al. 2014). The other classifiers such as Decision Trees, KNN and Ensemble are used to illustrate the effectiveness of the proposed classification framework.

SVMs are more appropriate (Burges 1998; Rouslan 2008) and can provide good generalization even if the training set has some bias; this gives unique solution because the loss function is convex. SVMs are nonparametric models whose complexity grows quadratically with the increase of number of record (Vempati et al. 2010). SVMs are even more suitable to small datasets with many features (Huang and LeCun 2006). As the large number of features results in curse of dimensionality which implies that to obtain the good generalization, the number of training samples must grow exponentially with the number of features (Huang and LeCun 2006; Übeyli 2010). Moreover, shallow architectures have practical limitations for efficient representation of certain types of function families (Bengio and Lecun 2007). To avoid these issues, it is required to generate such models that could capture the large degree of variation that can occurs in the underlying data pattern without having enumerate all of them. It is required to use the compact representation of data to capture most variation to reduce the curse of dimensionality as well as to reduce the computational complexity (Huang and LeCun 2006; Bengio and Lecun 2007; Erfani et al. 2015).

Support vector machine (SVM)

For supervised learning methods, SVM is one of the most robust method used for classification purposes. Recently, SVM is excellently used for pattern recognition problems (Vapnik 1999), machine learning (Gammerman et al. 2016) and medical diagnosis area (Dobrowolski et al. 2012; Subasi 2013). Moreover, SVM is used in variety of applications such as recognition and detection, text recognition, content based image retrial, biometrics, speech recognition etc. SVM construct a hyperplane or set of hyperplanes in infinite or high dimensional space which can be used for classification a good separation using this hyperplane is achieved that has the largest distance to the nearest training data point of any class (also known as functional margin), generally larger the margin indicates the lower generalization error of the classifier. SVM tries to find a hyperplane that gives the largest minimum distance to the training example. In SVM theory this name is also known as margin. For maximized hyperplane, the optimal margin is obtained. SVM has another important characteristic that gives the greater generalization performance. SVM is basically, a two-category classifier which transformed data into a hyperplane depends on the nonlinear training data or higher dimension.

Let us define a hyperplane by x.w + b = 0, where w is its normal. The linearly separable data is labelled as:

$$\{x_i, y_i\}, x_i \in R^N d, y_i \in \{-1, 1\}, \quad i = 1, 2, \ldots, N \quad (11)$$

Here y_i is the class label of two class SVM. By minimizing the objective function, the optimum boundary is obtained with maximal margin i.e. $E = w^2$ subject to

$$x_i \cdot w + b \geq 1 \, for \; y_i = +1$$

$$x_i \cdot w + b \leq 1 \; for \; y_i = -1 \tag{12}$$

Combining these into set of inequalities as

$$(x_i \cdot b + b)y_i \geq 1 \; for \; all \; i$$

Generally, the data is not linearly separable, in such cases a slack variable Ξ_i is used to denote the amount of misclassification rate. Thus, new subjective function is then reformulated as (Fig. 2):

$$E = \frac{1}{2}w^2 + C\sum_i L(\Xi_i) \tag{13}$$

Subject to

$$(x_i \cdot b + b)y_i \geq 1 - \xi_i \; for \; all \; i$$

Here first term on right hand side is regularization term that gives SVM an ability to generalize well on sparse data. The second term denote the empirical risk that those points which are misclassified or lie within the margin. L denote the cost function and C denote the hyper parameter representing trad-off effect by minimizing the empirical risk against maximizing the margin. Linear-error cost function is most commonly used because of its ability to detect the outliers. The dual formulation with $L(\Xi_i) = \Xi_i is$

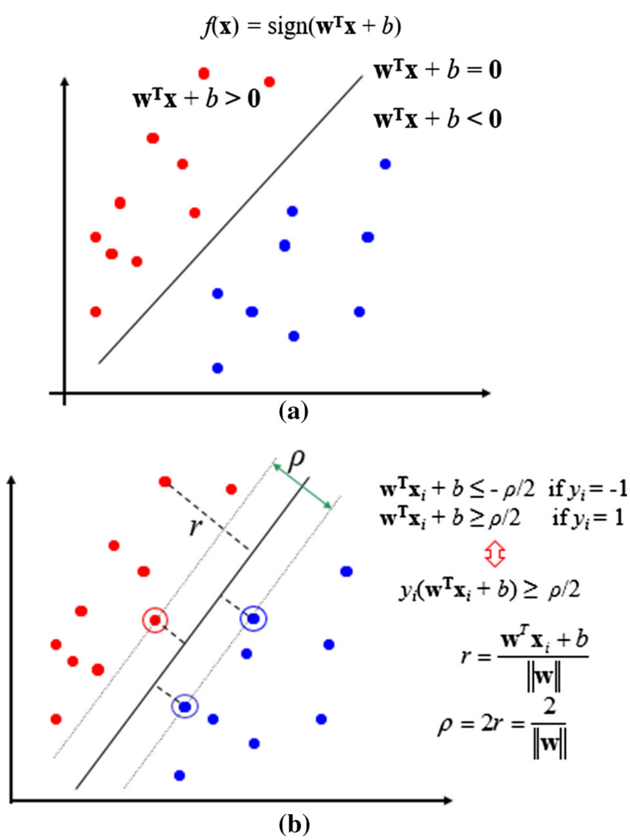

Fig. 2 SVM. **a** Linear separation and **b** margin

$$\alpha^* = max_\alpha\left(\sum_i \alpha_i + \sum_{i,j} \alpha_i\alpha_j y_i y_j x_i x_j\right) \tag{14}$$

Subject to

$$0 \leq \alpha_i \leq C and \sum_i \alpha_i y_j = 0$$

In which $\alpha = \{\alpha_1, \alpha_2, \alpha_3, \ldots, \alpha_i,\}$ is a set of Lagrange multipliers of the constraints in the primal optimization problem. The optimal decision boundary is now given by.

$$w_0 = \sum_i \alpha_i x_i y_i \tag{15}$$

SVM for non-linearly separable data

Muller et al. (2001) recommended the use of kernel function trick to deal with the data which is not linearly separable. In this case the non-linear mapping from input space is made to higher dimensional feature space. The dot product between two vectors in the input space is expressed by dot product with some kernel functions in the feature space. The most commonly used kernel functions are polynomial and radial base function (RBF). Mathematically, these are expressed as:

Types of Different Machine Learning Kernels with formulae.

SVM Polynomial Kernel

$$K(x_i, y_i) = (x_i \cdot y_i + 1)^n \tag{16}$$

SVM Gaussian (RBF) kernel

$$K(x_i, y_i) = exp\left(\frac{-1}{2}\frac{x_i - y_i^2}{\sigma^2}\right) \tag{17}$$

SVM Fine Gaussian (RBF) kernel

$$K(x_i, y_i) = exp\left(\frac{-1}{2}\frac{x_i - y_i^{'}||x_i - y_i||}{\sigma^2}\right) \tag{18}$$

Where n is the order of polynomial kernel and σ is the width of RBF. The dual formulation for non-linear case is given by

$$\alpha^* = max_\alpha\left(\sum_i \alpha_i + \sum_{i,j} \alpha_i\alpha_j y_i y_j K\left(x_i.x_j\right)\right) \tag{19}$$

Subject to

$$0 \leq \alpha_i \leq C \; and \; \sum_i \alpha_i y_j = 0$$

The SVM classifier performance depends on several parameters. The grid search method (Huang and LeCun 2006) was used to select the optimal parameter value by carefully setting grid range and step size. The linear kernel involves only one parameter ('c' soft margin constant), that represent the constraint violation cost associated with the data point occurring on the wrong side of the decision surface. The SVM with RBF and Gaussian kernel function has two training parameters: cost (C) which control the overfitting of the model and sigma (), which control the degree of nonlinearity of the model. The default values of cost function and sigma are used.

Decision tree (DTs)

In Decision Tree, the classifier check the similarities in the dataset and classify it into distinct classes. Wang et al. (2015) used DTs for classifying the data based on choice of an attribute which maximizes and fix the data division. Until the termination criteria is met, the attributes are split into several branches. Mathematically, a DT algorithm is constructed using following subsequent equations.

$$\overline{X} = \{X_1, X_2, X_3, \ldots, X_m\}^T \tag{20}$$

$$X_i = \{x_1, x_2, x_3, \ldots, x_{ij}, \ldots, x_{in}\} \tag{21}$$

$$S = \{S_1, S_2, , \ldots, S_i, \ldots .., S_m\} \tag{22}$$

where m represents the available observations number, n denote the independent variable number, S us the m-dimension vector of the variable forecasted from $\overline{X}$. X_i is the ith component of n-dimension autonomous variables $x_{i1}, x_{i2}, x_{i3}, \ldots, x_{in}$ are autonomous variable of pattern vector X_i and T is the transpose notation.

The purpose of DTs is to forecast the observations of $\overline{X}$. From $\overline{X}$ several DTs can be built with different accuracy level; however, an optimal DT is challenging because search space has large dimension. For DT suitable algorithms can be developed to reflect the trade-off between accuracy and complexity. In this case a sequence of local optimal decisions about the feature parameters are used to partition the dataset $\overline{X}$ using DT algorithms. Optimal DT, T_{k0} is constructed according to the subsequent optimization problem.

$$\hat{R}(T_{k0}) = min\{\hat{R}(T_{k0})\}, \quad k = 1, 2, 3, \ldots, K \tag{23}$$

$$\hat{R}(T) = \sum_{t \in T}^{k} \{r(t)p(t)\} \tag{24}$$

where $\hat{R}(T)$ denote the error level during the misclassification of tree T_k, T_{k0} denote the optimal DT that minimizes the error of misclassification in the binary tree, T denote the binary tree $\in \{T_1, T_2, , \ldots, T_k, t_1\}$, the index of tree is denoted by k, tree node by t, root node by t1, resubstituting error by r(t) that misclassify node t, probability that any case drop into node t is denoted by p(t). T^L *and* T^R denote the sub-trees of left and right partition set. The tree T is formed by feature plan portioning.

Ensemble classifiers

The ensemble classifiers comprise of set of individually trained classifiers whose predictions are then combined when classifying the novel instances using different approaches (Hussain et al. 2015). These are learning algorithms which construct a set of classifiers and then classify new data points by taking the weight of their predictions. These methods have successfully been used to enhance the prediction power in variety of applications such as predicting signal peptide (Chou and Shen 2007), for predicting protein subcellular location (Chou and Shen 2007), predicting subcellular location (Chou and Bin 2007) and enzyme subfamily prediction (Chou 2005). In many applications, the combined classification approaches give relative better performance than the individual classifier. Hayat and Khan (2012) reported that individual classifiers are diverse and can make different errors during classification, but during combing the classifiers the error is reduced because error produced by one classifier can be compensated by the other classifier.

K-nearest neighbor (KNN)

KNN is most widely used algorithm in the field of machine learning, pattern recognition and many other areas. Zhang et al. (2011) used KNN for classification problems. This algorithm is also known as instance based (lazy learning) algorithm. A model or classifier is not immediately build but all training data samples are saved and waited until new observations need to be classified. This characteristic of lazy learning algorithm makes it better than eager learning, that construct classifier before new observation needs to be classified. Schwenker and Trentin (2014) investigated that this algorithm is also more significant when dynamic data is required to be changed and updated more rapidly. KNN with different distance metrics were employed. KNN algorithm works according to the following steps using Euclidean distance formula.

Step I: To train the system, provide the feature space to KNN

Step II: Measure distance using Euclidean distance formula

$$d(x_i, y_i) = \sum_{i=1}^{n} \sqrt{(x_{i-}y_i)^2} \tag{25}$$

Step III: Sort the values calculated using Euclidean distance using $d_i \leq d_i + 1,$ *where* $i = 1, 2, 3, \ldots, k$

Step IV: Apply means or voting according to the nature of data

Step V: Value of K (i.e. number of nearest Neighbors) depends upon the volume and nature of data provided to

KNN. For large data, the value of k is kept as large, whereas for small data the value of k is also kept small.

Training/testing data formulation

The Jack–Knife tenfold cross validation technique was applied for training/testing data formulation and parameter optimization. It is one of the most well-known and commonly practiced and successfully used method to validate the accuracy of a classifier using tenfold CV, the data is divided into tenfolds, in training, the ninefolds participate and classes of samples of remaining folds are predicted based on the training performed on ninefolds. For the trained models, the test samples in test fold are purely unseen. The entire process is repeated 10 times and each class sample is predicted accordingly. Finally, the unseen samples predicted labels are used to determine the classification accuracy. This process is repeated for each combination of system's parameters, and classification performance have been reported for the samples as depicted in the Tables 1, 2, 3, 4, 5 and 6.

Results

The classification performance in form of TPR, TNR, PPV, NPV, overall accuracy and AUC was evaluated using most robust machine learning classifiers such as Support Vector Machine Kernels, Decision Trees, KNNs and Ensemble methods. The classifiers are tuned with different options and parameters to investigate the performance in more deeper detail than the ordinary classification methods used in the research. This gives an insight to see the performance of different classifiers in multiple scales, distance metrics, learning rates and distance weights etc. Feature extraction and selection and use of classifier is most important factor for proper analysis of any problem. In the past, researchers extracted different features for detection of epileptic seizure. In this study, we used different features extracting strategies keeping in view the dynamical properties of highly complex nonlinear, spatiotemporal variations of brain signals. Thus, we extracted the time and frequency domain features, complexity base features, wavelet entropy features and few statistical features from both healthy and epileptic and postical subjects. The most novel machine learning classifiers such as Support Vector Machine, Decision Trees, KNN and Ensemble classifiers

Table 1 Confusion matrix and comparison of performance evaluation

	Predicted Class			
Actual Class	**Support Vector Machine with Linear Kernel**			
	True Positive (TP) = 99	False Positive (FP) = 2	PPV=TP/(TP+FP) =99 / (99+2) =98%	TPR=TP/(TP+FN) = 99/ (99+0) = 100%
	False Negative (FN) = 0	True Negative (TN) = 99	NPV=TN/(TN+FN) = 99/ (99+0) =100 %	TNR=TN/(TN+FP) = 99/ (99+2) =98%
	Cosine KNN			
	True Positive (TP) = 95	False Positive (FP) = 6	PPV=TP/(TP+FP) =95 / (95+6) =94.1%	TPR=TP/(TP+FN) = 95/ (95+0) = 100%
	False Negative (FN) = 0	True Negative (TN) = 99	NPV=FP/(TP+FP) = 6/ (95+6) =100 %	TNR=TN/(TN+FP) = 99/ (99+6) 94.3%
	Coarse KNN			
	True Positive (TP) = 74	False Positive (FP) = 27	PPV=TP/(TP+FP) =74 / (74+27) =73.2%	TPR=TP/(TP+FN) = 74/ (74+0) = 100%
	False Negative (FN) = 0	True Negative (TN) = 99	NPV=TN/(TN+FN) = 99/ (99+0) =100 %	TNR=TN/(TN+FP) = 99/ (99+27) 78.5%

Table 2 Classification results to distinguish healthy set O subjects from epileptic seizure (set S, ictal intervals) using different machine learning classifiers

Classifier	TNR (%)	TPR (%)	NPV (%)	PPV (%)	Overall accuracy (%)	AUC	95% CI	
							LB	UB
Support vector machine (SVM)								
Linear	98.0	100	100	98.0	99.0	0.9991	0.00	0.98
Quadratic	97.0	100	100	97.1	98.5	0.9887	0.00	0.97
Fine Gaussian	99.0	98.0	98.0	99.0	98.5	0.9922	0.03	0.99
Coarse Gaussian	94.1	100	100	94.3	97.0	0.9986	0.00	0.93
Nearest neighbor								
Fine KNN	97.0	98.0	98.0	97.0	97.5	0.9750	0.01	0.98
Coarse KNN	74.3	100	100	79.2	87.0	0.9961	0.00	0.75
Weighted KNN	97.0	99.0	99.0	97.0	98.0	0.9940	0.00	0.99
Decision tree								
Complex tree	99.0	98.0	98.0	99.0	98.5	0.9815	0.02	0.99
Medium tree	99.0	98.0	98.0	99.0	98.5	0.9815	0.02	0.99
Ensemble								
Boosted tree	100	88.9	90.2	100	94.5	0.9001	0.00	0.96
Bagged tree	98.0	99.0	99.0	98.0	98.5	0.9996	0.01	0.98
Subspace disc.	95.0	100	100	95.2	97.5	0.9986	0.00	0.94
Subspace KNN	99.0	99.0	99.0	98.0	98.5	0.9992	0.01	0.98

Table 3 Performance Evaluation Results to Distinguish Healthy Set O subjects from epileptic seizure (Set S, ictal intervals) using support vector machine (SVM) classifier with linear kernel by tuning parameters with fixed box constraint level (BCL) value of 1 and different range of values for kernel scale (KS) and vice versa

KS	TNR (%)	TPR (%)	NPV (%)	PPV (%)	Overall accuracy (%)	AUC
1	99.0	100	100	99.0	99.5	0.9987
2–6	98.0	100	100	98.0	99.0	0.9987
7–9	97.0	100	100	97.1	98.5	0.9989
10–11	96.0	100	100	96.1	98.0	0.9988
12	95.0	100	100	95.2	95.7	0.9991
13–15	93.1	100	100	93.4	96.5	0.9990
16	92.1	100	100	92.5	96.0	0.9989
17	93.1	100	100	93.4	96.5	0.9988
18–19	92.1	100	100	92.5	96.0	0.9985
20	90.1	100	100	90.8	95.0	0.9984
25	84.2	100	100	86.1	92.0	0.9974
30	78.2	100	100	81.8	89.0	0.9977
40	64.4	100	100	73.3	82.0	0.9982
50	55.4	100	100	68.8	77.5	0.9981
BCL	**TNR (%)**	**TPR (%)**	**NPV (%)**	**PPV (%)**	**Overall accuracy (%)**	**AUC**
2–19	99.0	100	100	99.0	99.5	0.9986
20–23	98.0	100	100	98.0	99.0	0.9990
> 24	97.0	100	100	97.1	98.5	0.9982

with distinct set of possible learner parameters are used to improve the evaluation performance. This will give a new direction to judge the performance on broader parameters selection rather than choosing few options. Figure 4 below shows the scatter plots again two selected features such as SDSD along x-axis and Entropy values on y-axis to

Table 4 Performance evaluation results to distinguish healthy set O subjects from epileptic seizure (set S, ictal intervals) using support vector machine (SVM) classifier with quadratic kernel by tuning parameters with fixed box constraint level (BCL) value of 1 and different range of values for kernel scale (KS) and vice versa

KS	TNR (%)	TPR (%)	NPV (%)	PPV (%)	Overall accuracy (%)	AUC
1	96.0	99.0	99.0	96.1	97.5	0.9982
2–3	97.0	100	100	97.1	98.5	0.9986
4–9	98.0	100	100	98.0	99.0	0.9989
10–14	97.0	100	100	97.1	98.5	0.9989
15–16	96.0	100	100	96.1	98.0	0.9989
17	95.0	100	100	95.2	97.5	0.9991
18	94.1	100	100	94.3	97.0	0.9990
19–24	93.1	100	100	93.4	96.5	0.9991
25	92.1	100	100	92.5	96.0	0.9986
30	89.1	100	100	90.0	94.5	0.9984
40	81.2	100	100	83.9	90.5	0.9978
50	71.3	100	100	77.3	85.5	0.9977
BCL	**TNR (%)**	**TPR (%)**	**NPV (%)**	**PPV (%)**	**Overall accuracy (%)**	**AUC**
> 2	95.0	98.0	98.0	95.1	96.5	0.9870

Table 5 Performance evaluation results to distinguish healthy set O subjects from epileptic seizure (set S, ictal intervals) using support vector machine (SVM) classifier with Gaussian Kernel by tuning parameters with fixed box constraint level (BCL) value of 1 and different range of values for kernel scale (KS) and vice versa

KS	TNR (%)	TPR (%)	NPV (%)	PPV (%)	Overall accuracy (%)	AUC
1	99.0	94.9	95.2	98.9	97.0	0.9946
2–14	98.0	99.0	99.0	98.0	98.5	0.9964
15	96.0	100	100	96.1	98.0	0.9990
16	95.0	100	100	95.2	97.5	0.9981
17	94.1	100	100	94.3	97.0	0.9991
18–24	93.1	100	100	93.4	96.5	0.9991
25	92.1	100	100	92.5	96.0	0.9985
30	89.1	100	100	90.0	94.5	0.9982
40	81.2	100	100	83.9	90.5	0.9976
50	71.3	100	100	73.3	85.5	0.9977
BCL	**TNR (%)**	**TPR (%)**	**NPV (%)**	**PPV (%)**	**Overall accuracy (%)**	**AUC**
> 2	99.0	96.0	96.2	99.0	97.5	0.9943

Table 6 Performance evaluation results to distinguish healthy set O subjects from epileptic seizure (set S, ictal intervals) using K-nearest neighbor classifier with varying distance metrics and selection criterion of number of neighbors as 1 and equal distance weight

Distance metric	TNR (%)	TPR (%)	NPV (%)	PPV (%)	Overall accuracy (%)	AUC
Euclidean distance	98.0	99.0	99.0	98.0	98.5	0.9850
City block	99.0	100	100	99.0	99.5	0.9950
Chebyshev	95.0	99.0	99.0	95.1	97.0	0.9702
Cubic	93.1	100	100	93.4	96.5	0.9653
Cosine	97.0	99.0	99.0	97.0	98.0	0.9800
Correlation	97.0	99.0	99.0	97.0	98.0	0.9800
Spearman	98.0	99.0	99.0	98.0	98.5	0.9850
Hamming	54.5	97.0	94.0	67.6	75.5	0.7512
Jaccard	54.5	97.0	94.0	67.6	75.5	0.7512

distinguish the health and epileptic subjects. For the illustration purposes, we selected SVM linear kernel, cosine KNN and coarse KNN. The left column figures denote the true class while right column figures denote the predicted

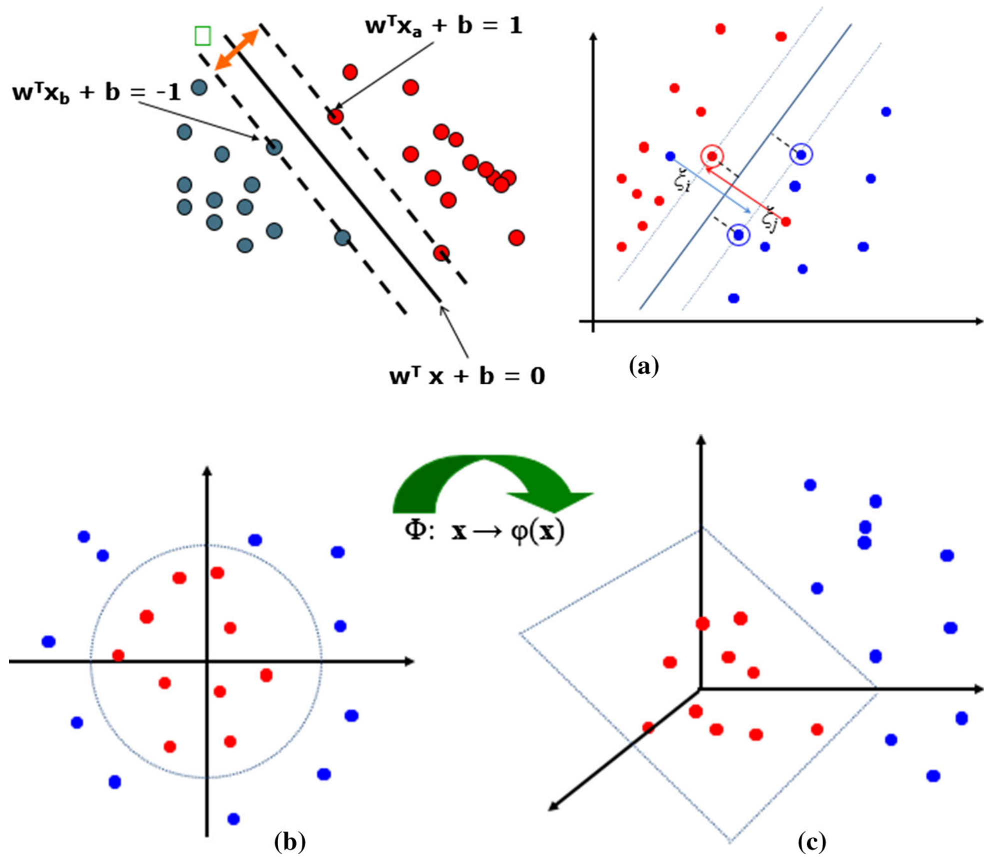

Fig. 3 **a** Error on margin using slack variable, **b**, **c** SVM non-linear separation

class. The red color dots denote the healthy and green color dots denote the epileptic subjects. Moreover, cross in both colors represents the errors. The correspondence confusion matrix clearly shows the performance measures for these three classifiers. There were 2, 6 and 27 data points which are predicted as false positive after prediction. The scatter plots also reflect the red crosses for true class and green cross for predicted class. Thus, scatter plots help us to see the classification performance visually for true classes and predicted classes. A pictorial representation of error using slack variable and non-linear separation using SVM is shown in Fig. 3.

Receiver operating curve (ROC)

The ROC is plotted against the true positive rate (TPR) i.e. sensitivity and false positive rate (FPR) i.e. specificity values of healthy and epileptic seizure subjects. The mean features values for healthy subjects are classified as 1 and epileptic subjects are classified as 0. This vector is then passed to the ROC function, which plots each sample values against specificity and sensitivity values. ROC is a standard way to classify the performance and visualize the behavior of a diagnostic system (Hajian-Tilaki 2013). The TPR is plotted against y-axis and FPR is plotted against x-axis. The area under the curve (AUC) shows the portion of a square unit. Its value lies between 0 and 1. AUC > 0.5 shows the separation. The higher AUC shows the better diagnostic system. Correct positive cases divided by the total number of positive cases are represented by TPR, while negative cases predicted as positive divided by the total number of negative cases are represented by FPR. The SVM linear classifier shows separation as 0.99 and Coarse SVM as 0.75 as shown in the Fig. 6. The AUC values using different classifiers are depicted in Tables 1, 2, 3, 4, 5 and 6. The Fig. 5 shows parallel coordinates plot of data with predictions for SVM linear kernel and Coarse KNN classifiers against $means \pm n\ SD$, where $n = \{\pm\ 2, 4, 6, 8\}$ for different extracted features.

Performance measures based on confusion matrix parameters

To detect the epileptic seizure, following measures were used to compute the True Positive Rate (TPR), True Negative Rate (TNR), Positive Predictive Value (PPV), Negative Predictive Value (NPV), Overall Accuracy and Area under receiver Operating Curve (AUC) with example of healthy Set O vs Epileptic Set S subjects as illustrated in Figures ROCs, Scatter plots and Confusion Matrix and performance evaluation Tables below.

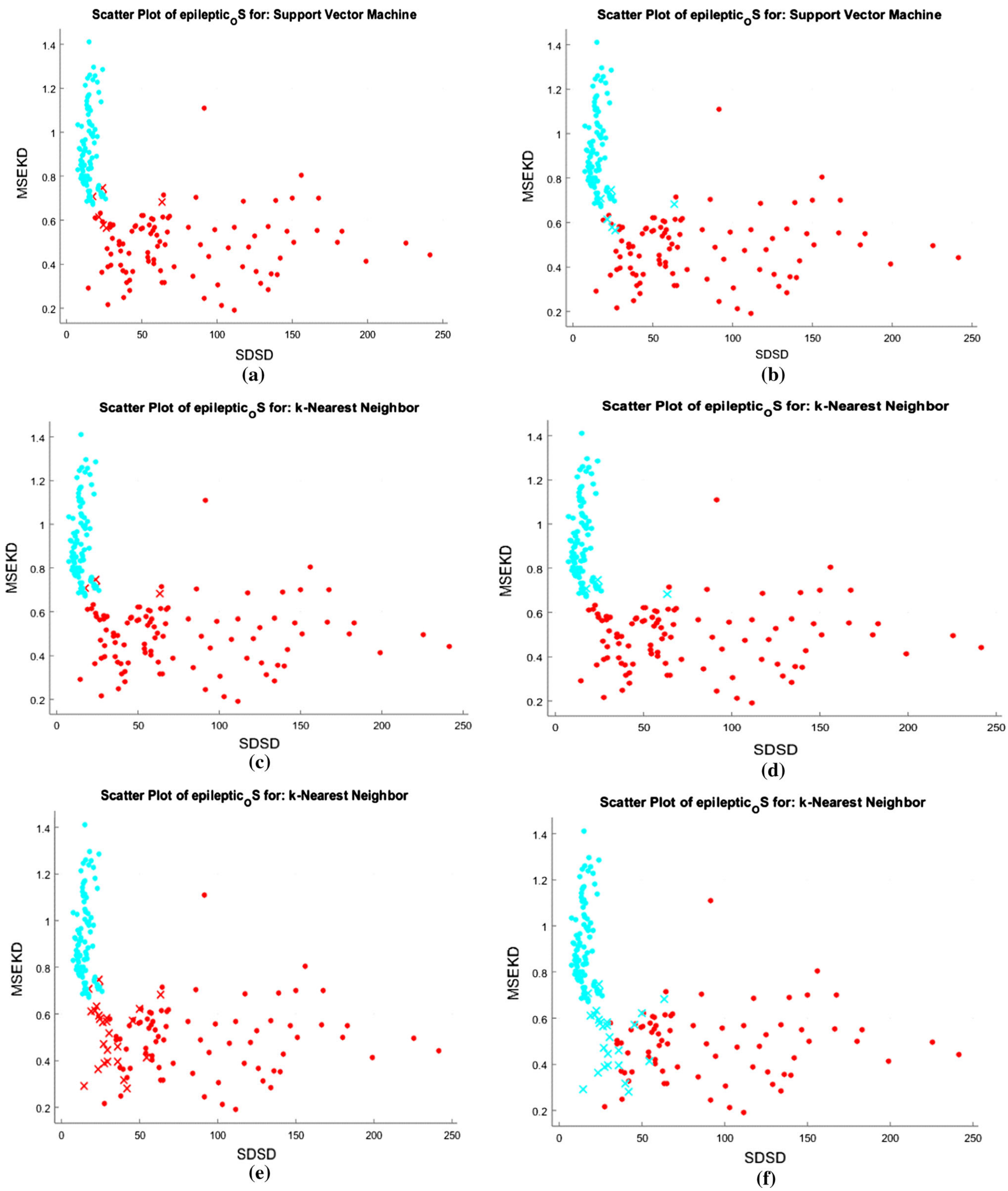

Fig. 4 **a**, **b** SVM linear kernel (left) true class (right) predicted class, **c**, **d** using cosine KNN (left) true class (right) predicted class, **e**, **f** using Coarse KNN (left) true class (right) predicted class

True positive rate (TPR)

The TPR measure also known as sensitivity or recall is used to test the proportion of people who test positive for the disease among those who have the disease. Mathematically, it is expressed as:

$$TPR = \frac{\sum True\ Positive}{\sum Condition\ Positive}$$

$$TPR = \frac{TP}{TP + FN}$$

i.e. the probability of positive test given that patient has disease.

True negative rate (TNR)

The TNR measure also known as Specificity is the proportion of negatives that are correctly identified. Mathematically, it is expressed as:

$$TNR = \frac{\sum True\ Negative}{\sum Condition\ Negative}$$

$$TNR = \frac{TN}{TN + FP}$$

i.e. probability of a negative test given that patient is well.

Positive predictive value (PPV)

PPV is mathematically expressed as:

$$PPV = \frac{\sum True\ Positive}{\sum Predicted\ Condition\ Positive}$$

$$PPV = \frac{TP}{TP + FP}$$

where TP denote that the test makes a positive prediction and subject has a positive result under gold standard while FP is the event that test make a positive perdition and subject make a negative result.

Negative predictive value (NPV)

NPV can be computed as:

$$NPV = \frac{\sum True\ Negative}{\sum Predicted\ Condition\ Negative}$$

$$NPV = \frac{TN}{TN + FN}$$

where TN indicates that test make negative prediction and subject has also negative result, while FN indicate that test make negative prediction and subject has positive result.

Overall accuracy (OA)

The total accuracy is computed as:

$$TA = \frac{TP + TN}{TP + FP + FN + TN}$$

Based on confusion matrix, using confusion matrix and scatter plot, PPV (98%) was obtained using SVM Linear kernel, while PPV (94.1%) using Cosine KNN and PPV (73.2%) was obtained using Coarse KNN. The scatterplots also show the same resemblance visually, that true class (left) and predicted class (right), after prediction, there were wo False positive data points using SVM Linear kernel, six False positive data points using Cosine KNN and 27 False positive data points using Coarse KNN. Thus, scatterplots also help us to see the relationships between actual and predicted classes to distinguish healthy and pathological conditions based on any feature values and shows the misclassification percentage for true and predicted classes (Figs. 5, 6, 7).

In Fig. 5, the blue color denotes the means of epileptic subjects (Set S) and red color denote the healthy (Set O) subjects. The lines denote the correctly classified subjects, while x denote the incorrectly classified samples using (a) SVM linear kernel and (b) Coarse KNN. It is evident from the Fig. 5a, b and results computed as reflected in Tables 1 and 2 that there are less incorrectly classified data points using SVM linear kernel, while Coarse KNN give higher misclassification result as depicted in Fig. 5b. The model also shows that healthy subjects give much higher incorrectly classification results than the epileptic subjects using Coarse KNN, while SVM linear kernel, there are fewer incorrectly classified samples. The Fig. 8a–d reflect the mean features values plotted against healthy and epileptic subjects using multi-features vector of (a) frequency domain, (b) time domain and statistical, (c) wavelet entropy and (d) complexity based fast sample entropy features. The results from Fig. 5a, b and a–d shows that epileptic mean features values of epileptic subjects are much greater than the healthy subjects for each case (a–c) because epileptic subjects are produced due to higher neurological chronic disorder and higher spikes are produced. The epileptic feature values of healthy subjects are higher than the epileptic subjects because healthy subjects exhibit higher complexity than the epileptic subjects which is also consistent with the previous studies (Hussain et al. 2017b; c; Costa et al. 2002).

The Table 2, depicts the classification performance evaluation results using SVM, KNN, Decision Tree and Ensemble classifiers at default parameters for each classifier. In this study, we employed novel ML classifiers such as Support Vector Machine (SVM) with Liner, Quadratic, Fine Gaussian, Coarse Gaussian; KNN with Fine KNN, Coarse

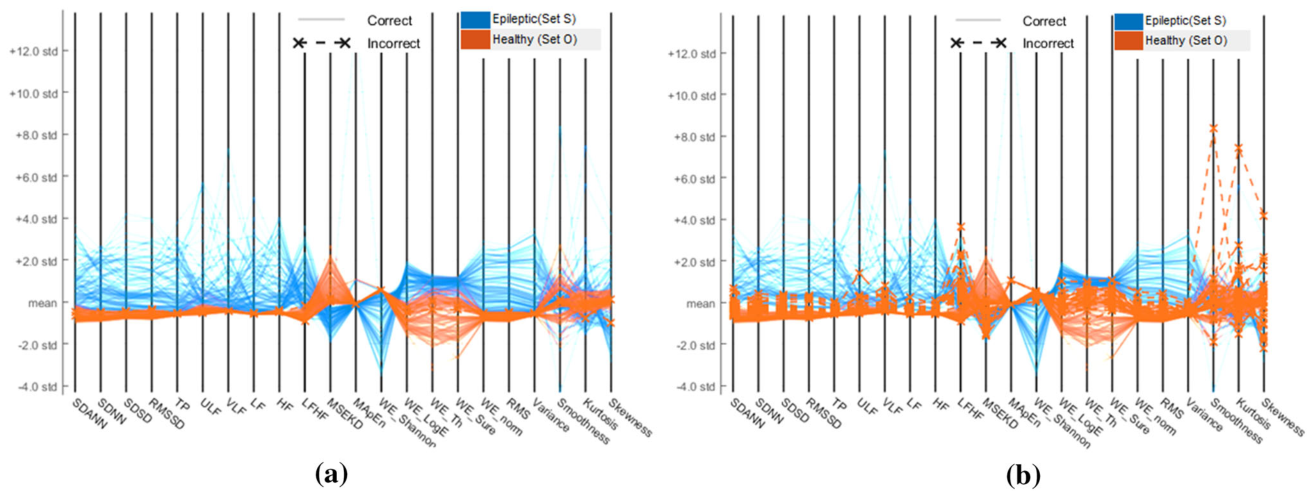

Fig. 5 Detection of Epileptic Seizure using varying feature set and classifying using **a** SVM linear kernel, **b** Coarse KNN

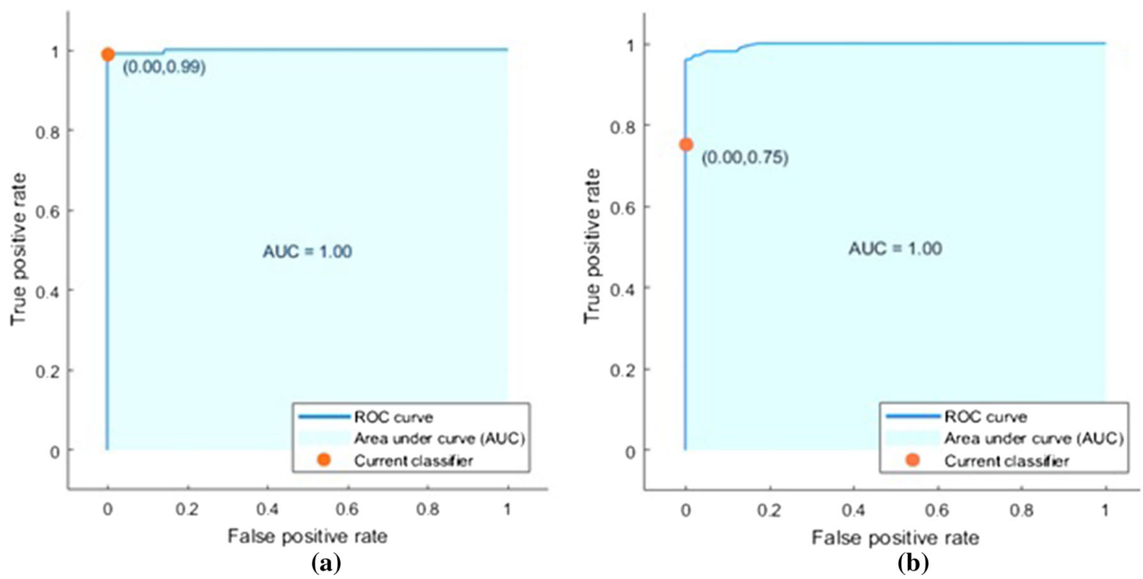

Fig. 6 Receiver Operating Curve (ROC) to distinguish healthy and epileptic subjects using **a** SVM linear kernel, **b** Coarse KNN

KNN, Weighted KNN; Decision Tree with Complex Tree and Medium Tree and Ensemble Classifiers such as Boosted Tree, Bagged Tree, Subspace Tree Discriminant and Subspace KNN. We extracted varying features extracting strategy and combined the feature vectors used as input to these classifiers. Using SVM, the highest TPR and NPV of 100% were obtained using Linear, Quadratic and Coarse Gaussian kernels followed by Fine Gaussian of 98%. Moreover, highest TNR was obtained as 99% with Fine Gaussian followed by Linear (98%), Quadratic (97%) and Coarse Gaussian (94.1%). Likewise, we obtained the highest PPV with Fine Gaussian (99%), followed by Linear (98%), Quadratic (97.1%), and Coarse Gaussian (94.3%). The overall accuracy using SVM was obtained with linear kernel (99%) followed by Quadratic and Fine Gaussian (98.5%) and Coarse Gaussian (97%). Similarly, the highest separation in form of AUC was obtained using Linear kernel ($AUC = 0.9991$) followed by Fine Gaussian ($AUC = 0.9922$), Coarse Gaussian ($AUC = 0.9987$) and Quadratic kernel ($AUC = 0.9887$). Similarly, using KNN, the highest accuracy was obtained with Weighted KNN (98%) followed by Fine KNN (97.5%), and Coarse KNN (87%). The highest separation using KNN was found using Coarse KNN ($AUC = 0.9961$) followed by Weighted KNN ($AUC = 0.9940$) and Fine KNN ($AUC = 0.9750$). The other performance evaluation values are reflected in Table 2. The Decision Tree with both Complex and Medium Trees give accuracy of 98.5% and AUC of 0.9815. Similarly, using Ensemble methods, the highest accuracy of 98.5% was obtained with Bagged Tree and Subspace KNN followed by

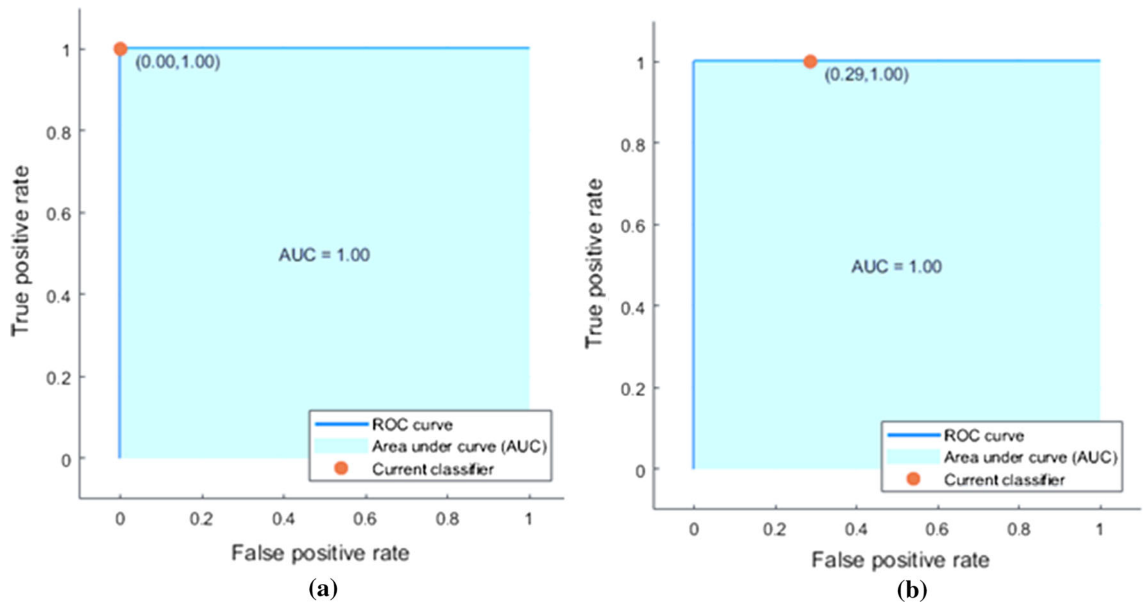

Fig. 7 Receiver Operating Curve (ROC) to distinguish Postictal heart rate oscillations from epileptic subjects using **a** SVM linear kernel, **b** weighted KNN

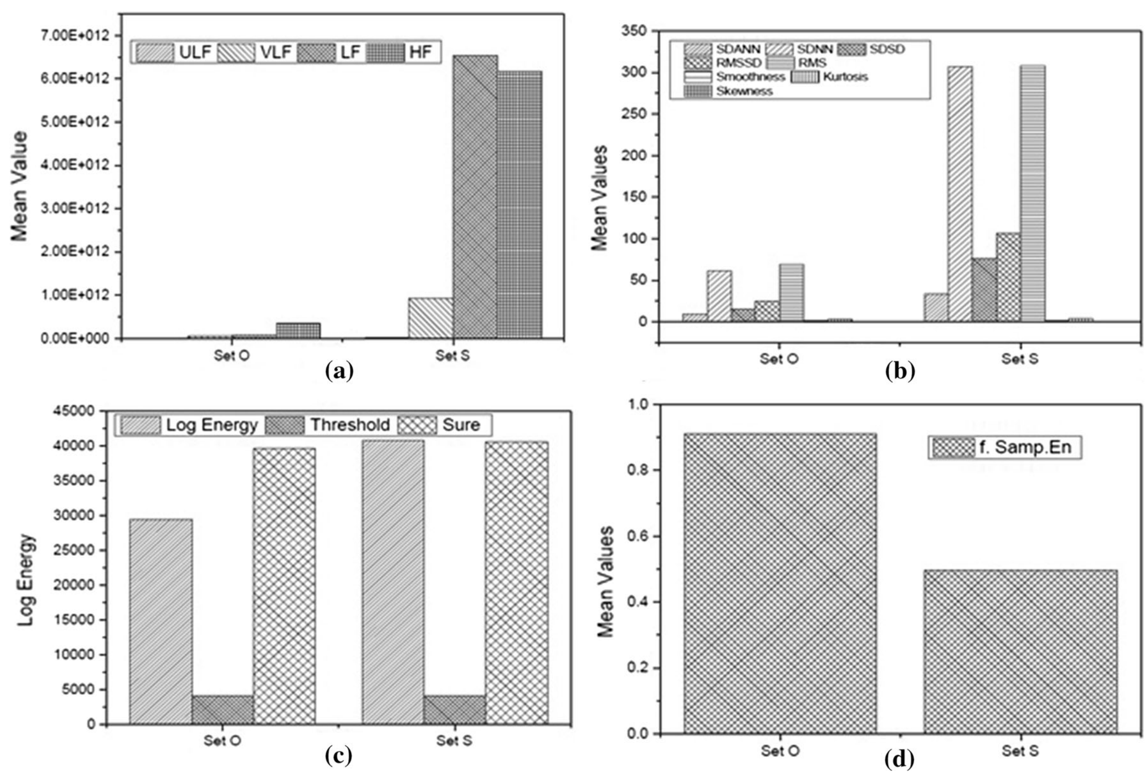

Fig. 8 Mean values of multidomain features extracted from healthy and epileptic subjects. **a** Frequency domain features, **b** time domain and statistical features, **c** wavelet entropy features, **d** complexity base features

Subspace Discriminant (97.5%) and Boosted Tree as 94.5%. AUC of 0.9996 was obtained with Bagged Tree followed by Subspace KNN (AUC = 0.9992), Subspace Discriminant (AUC = 0.9986) and Boosted Tree (AUC = 0.9001). The 95% confidence intervals about AUC are also depicted in the Tables 2 and 7.

Table 7 Classification Results to Distinguish Post ictal heart rate oscillations with partial epilepsy subjects from epileptic seizure (Set S, ictal intervals) using Different Machine Learning Classifiers

Classifier	TNR (%)	TPR (%)	NPV (%)	PPV (%)	Overall accuracy (%)	AUC	95% CI	
							LB	UB
Support vector machine (SVM)								
Linear	100	100	100	100	100	1.00	0.00	1.00
Quadratic	100	100	100	100	100	1.00	0.00	1.00
Fine Gaussian	100	97.09	54.14	100	97.2	1.00	0.43	1.00
Coarse Gaussian	100	99.01	85.71	100	99.1	1.00	0.14	1.00
Nearest neighbor								
Fine KNN	100	100	100	100	100	1.00	0.00	1.00
Coarse KNN	100	99.01	85.71	100	99.1	1.00	0.16	1.00
Weighted KNN	100	98.04	71.43	100	98.1	1.00	0.29	1.00
Decision tree								
Complex tree	100	100	100	100	100	1.00	0.00	1.00
Medium tree	100	100	100	100	100	1.00	0.00	1.00
Ensemble								
Bagged tree	100	100	100	100	100	1.00	0.00	1.00
Subspace disc.	100	100	100	100	100	1.00	0.00	1.00
Subspace KNN	100	100	100	100	100	1.00	0.00	1.00

The Table 3 depicts the performance evaluation using SVM linear kernel in more depth with multiple kernel scales and box constraint level. In the first case, we used multiple kernel scales and fixed box constraint value of 1. The performance was measured as reflected in the Table 3, it was seen that performance metrics such as TNR, PPV, overall accuracy and AUC vary at multiple kernel scales. This Kernel improved the overall accuracy from 99 to 99.5% with KS = 1 and BCL = 1, when we changed the default kernel scale and box constraint levels from the default values for Linear SVM. The highest accuracy at other kernel scales (KS 2–6) was obtained as 99%, followed by 98.5% (KS 7–9), 98% (KS 10–11), and so on. It decreases as we increased the KS and obtained 77.5% at KS 55. However, TPR, NPV remained the same as 100%. The TNR and PPV values have similarly performance as obtained for overall accuracy. Likewise, maximum separation was obtained at KS 12 of (AUC = 0.9991) followed by KS 13–15 (AUC = 0.9990), KS 7–9,16 (AUC = 0.9989), KS 10–11, 17 (AUC = 0.9988), KS 1, 2–6 (AUC = 0.9987), KS 18–19 (AUC = 0.9985) and decreased for other kernel scales. The AUC is also one of the most important factor to distinguish and get separation between the healthy and pathological subjects. The AUC values at different KS shows that there are more than one kernel scales are important for separating these subjects. Thus, these deeper parameters give more insight towards the improved performance. In the second case, we fixed the KS and check the evaluation performance with different BCL. It was observed that overall highest accuracy was obtained with BCL 2–19 as of 99.5% followed by BCL 20–23 of 99% and 98.5% for BCL > 24. The highest AUC was obtained at BCL 20–23 of AUC = 0.9990 followed by BCL 2–19 of 0.9986 etc.

This paper depicts the evaluation performance using SVM Quadratic kernel at multiple kernel scales and BCL to distinguish the healthy and epileptic subjects. Again here, the overall highest accuracy was obtained at KS 4–9 as of 99% improved than 88.5% with fixed scales and BCL followed by 98.5% at KS 2–3, 10–16, 97.5% at KS 1, 17 and so on. The highest separation was obtained at KS 17, 19–24 with (AUC = 0.9991) followed by KS 18 (AUC = 0.9990), KS 4–16 (AUC = 0.9989), KS 2–3, 25 (AUC = 0.9986) and so on. However, with fixed KS of 1 and BCL > 2, we obtained overall accuracy of 96.5%, PPV of 95.1%, TNR of 95%, TPR and NPV of 98%. And AUC of 0.9870.

The Table 5 shows the results using SVM with Gaussian kernel with varying KS and BCL values. In this case, we obtained an overall accuracy of 98.5% at KS 2–14, followed by 98% at KS 15, 97.5% at KS 16, 97% at KS 1 and 17 etc. Similarly, the highest AUC of 0.9991 was obtained at KS 17–24, AUC of 0.9990 at KS 15, AUC of 0.9985 at KS 25 and 0.9982 at KS 30. Accordingly, the TPR and PPV decreased as the KS increased, while TNR and NPV were found higher at all the KS. By fixing the KS as 1 and varying BCL > 2 values we obtained an overall accuracy of 97.5%, TNR 99%, TPR 96%, PPV 99%, NPV 96.2% and AUC = 0.9943.

In Table 6, we computed the evaluation performance using KNN with different distance metrics. The highest accuracy of 99.5% was obtained using KNN with distance metric of City Block followed by 98.5% with Euclidean Distance, Spearman; 98% using Cosine and Correlation; 97% using Chebyshev, 96.5% using Cubic, 75.5% using Hamming and Jaccard distance metric. Similarly, highest accuracy was obtained (AUC = 0.9950) using City Block, followed by (AUC = 0.9850) using Euclidean Distance & Spearman, AUC = 0.9800 using Cosine & Correlation, AUC = 0.9702 using Chebyshev, AUC = 0.9653 using Cubic and AUC = 0.7512 using Hamming and Jaccard distance metric. The other performance metrics such as TNR, TPR, NPV and PPV are reflected in Table 6.

To distinguish the postictal heart rate oscillations with partial epilepsy subjects from epileptic seizures, the support vector machine with linear and quadratic kernels, fine KNN, Ensemble methods including Bagged Tree, Subspace Discriminate, Subspace KNN give highest performance of 100% with respect to TPR, TNR, PPV, NPV and accuracy and AUC of 1.00. SVM fine Gaussian gives an accuracy of 99.1%, TNR (100%), TPR (99.01%), NPV (85.71%), PPV (100%), and AUC (1.00). SVM fine Gaussian gives an accuracy of 97.2%, TNR (100%), TPR (97.09%), NPV (54.14%), PPV (100%) and AUC (1.00). Likewise, a 100% of TNR and PPV were obtained using Fine, Coarse and Weighted KNN, while Coarse KNN gives TPR (99.01%), NPV (85.71%), and accuracy of 99.1% where weighted KNN gives TPR (98.04%), NPV (71.43%) and an accuracy (98.1%). The 95% confidence interval against each classifier is also reflected in Table 7. The comparison of classification performance is reflected in Table 8.

The Fig. 9 depicts the overall accuracy using KNN with varying Neighbors and distance weight. All the distance weights i.e. equal, inverse and squared inverse at smaller number of Neighbors i.e. 1 and 4 gives the highest accuracy of 98.5%, however, this accuracy decreases as we increased the Neighbors as reflected in the Fig. 9. Using the equal distance weight, the accuracy obtained with different Neighbors i.e. N8 (97.5%), N10 (97%), N20 (96%), N30 (96%), N50 (93.5%) and N100 (86.5%) N100. Likewise, using inverse distance weight, the accuracy obtained was N8 (98%), N10 (97.5%), N20 (96.5%), N30 (96%), N50 (95%) and N100 (90%). Similarly, using squared inverse distance weight, we obtained an accuracy as N8 (98.5%), N10 (98.5%), N20 (96.5%), N30 (96.5%), N50 (96%) and N100 (96%).

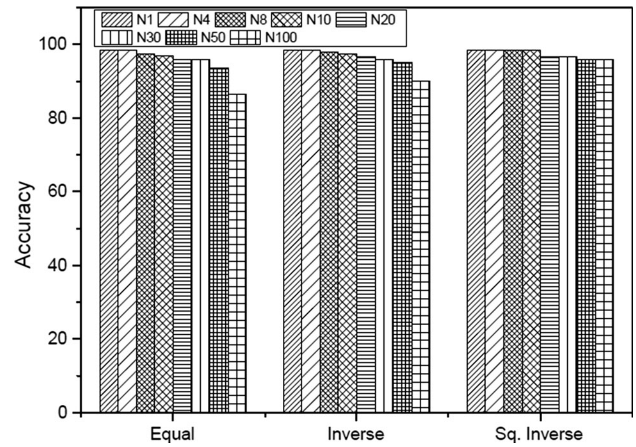

Fig. 9 Overall Accuracy obtained to Distinguish Healthy Set O subjects from epileptic seizure (Set S, ictal intervals) using K-Nearest Neighbor Classifier using Euclidean Distance with varying Neighbors (1, 4, 8, 10, 20, 30, 50 and 100) and distance weight (i.e. Equal, Inverse and Squared Inverse)

Discussions

In the past researchers used many complexity base measures to quantify the dynamics of physiological systems. Hussain et al. (2017c) recently used Symbolic time series

Table 8 The comparison of classification performance from different methods for the same data set

Problems	References	Methods	Accuracy (%)
S-FNOZ	Tzalla et al. (2007)	Time–frequency analysis, artificial neural network	97.73
	Guo et al. (2010)	Multiwavelet transform, MLPNN	98.27
	Rivero et al. (2011)	Time frequency analysis, KNN	98.40
	Kaleem et al. (2013)	Variation of empirical mode decomposition	98.20
	Fu et al. (2015)	HMS analysis, SVM	98.80
	Niknazar et al. (2013)	Wavelet transform, RQA, ECOC	98.67
	Peker et al. (2016)	Dual-tree complex wavelet transform, complex-valued neural networks	99.15
	Jaiswal and Banka (2017)	Local neighbor descriptive pattern, artificial neural network	98.72
	Wang et al. (2017)	DWT, multi-domain feature extraction and nonlinear analysis	99.25
	This work	Multidomain features, Complexity features based on entropy and KD tree algorithm, wavelet and statistical features	99.50

to detect and quantify the dynamics of epileptic seizure and distinguished between the healthy and epileptic subjects (ictal intervals) and interictal intervals (i.e. focal and non-focal signals) and compared the results with multiscale sample entropy. They found that healthy subjects have higher complexity than the epileptic subjects (with both ictal and interictal intervals). The pathological subjects lose their robustness due to degradation of structural components and coupling functions. Moreover, Hussain and Aziz (2016) used Time–Frequency Wavelet Phase Coherence to distinguish the EC from EO condition using multiple spatiotemporal scale. Recently, Hussain et al. employed KD tree algorithm based entropy and compared the results with traditional sample entropy. They observed that KD tree entropy is most effectively distinguishing the healthy and diseased subjects and more robust with respect to speed and memory. Thus, here we extracted complexity features based on KD tree algorithmic approach as described in detail in Hussain et al. (2017b).

Researchers extracted features from epileptic subjects using different strategies. Hassan and Bhuiyan (2016, Hassan and Subasi (2016) decomposed EEG signals into empirical mode decomposition with adaptive noise (CEEMDAN) and extracted six spectral movements namely spectral roll-off, decrease, centroid, spread, flatness and slope from CEEMDAN mode function. The classification was performed using ensemble machine learning linear programming boosting (LPBoost). These features selected are merely based on spectral components. The features extracted by analyzing the EEG signals solely in time domain may omit the important frequency information and vice versa. Moreover, EEG signals are extremely nonstationary and nonlinear and complex in nature which also requires extracting the features from both linear and nonlinear analysis methods. More recently, Wang et al. (2017) in 2017 detected epileptic seizure using multi-domain features extraction along with nonlinear analysis and obtained overall accuracy of 99.25% using Support Vector Machine (SVM) with default parameters. Researchers (Subasi et al. 2017; Subasi 2007; Polat and Güneş 2007; Acharya et al. 2015; Samiee et al. 2015; Kaya et al. 2014; Guo et al. 2009; Subasi and Erçelebi 2005) recently also evaluated the performance in terms of specificity, sensitivity and accuracy only with default parameters. In this study, the performance was evaluated in terms of TPR, NPR, PPV, NPV, accuracy and AUC by employing robust machine learning classification methods with detailed parametric analysis to distinguish the healthy and epileptic subjects using multimodal features extracting approach and obtained an improved accuracy of 99.50% with other performance evaluation parameters because based on few evaluation parameters performance cannot be fully judged for automated detection of epileptic seizure. The results are consistent with the previous studies using same database and features extracting strategy with an improved and detailed analysis performance. Moreover, prediction performance in term of individual features was also acquired as reflected in the prediction models in Fig. 5a, b and scatter plots Fig. 4a–f.

In this study, we proposed multidomain and nonlinear features extracting strategies to acquire the more effective approach for detecting the epileptic seizure with healthy subjects. We employed robust machine learning classifiers with advanced parametrization approach for better performance with overall accuracy of 99.50% with SVM and KNN using City Block Distance metric. Moreover, highest separation of (AUC = 0.9991) was obtained using these classifiers. The SVM kernels performance was evaluated on multiple scales. Likewise, KNN algorithms are computed based on different distance weights such as equal, square and inverse squared and number of Neighbors selected each time. We also used Jack-knife tenfold cross validation technique for training/testing purposes the most successful and commonly practiced techniques. The Table 2, give the performance results using different classifiers at default parameters. The other tables depict the evaluation performance in more deeper details using optimized parametrization approach. The Decision trees and Ensemble methods does not provide the performance by optimizing the parameters with different number of learners, learning rates, maximum number of splits and split criteria, but the results with these parameters does not changed may be due to the limitations of smaller database. Researchers in the past employed different approaches for features extracting approaches and classifiers to distinguish these subjects, such as average accuracy of 97.72% was acquired by (Tzallas et al. 2007) using time–frequency analysis, 98.27% by Guo et al. (2010) using multi-wavelet transform, 98.20 and 98.20% by Fu et al. (2015), Kaleem et al. (2013) using nonlinear analysis of empirical model decomposition, 99.15% by Peker et al. (2016) using complex valued classifier. In this work, we used multidomain features extracting strategy with time domain, frequency domain, statistical, nonlinear entropy base with KD tree algorithm and Wavelet entropy features using most robust machine learning learners with optimized parametric approach and obtained an overall accuracy (99.50%), TNR (99.0%), TPR (100%), NPV (100%), PPV (99.0%) and AUC of 0.9991 using tenfold cross validation. Moreover, this approach gives more effective and detailed analysis at multiple kernel scales, learning rates, and distance metrics and weights.

The results reveal that with default parameters and different features extracting strategies, the overall accuracy was obtained using Linear SVM of 99%, followed by Quadratic SVM, Fine Gaussian SVM, Complex & Medium

Tree, Bagged Tree, Subspace KNN of 98.5% followed by Fine KKN, Subspace Discriminant as 97.5%, and Coase Gaussian as 97%. The highest AUC was obtained using Bagged Tree (AUC = 0.9996) followed by Linear SVM (AUC = 0.9991), Subspace KNN (AUC = 0.9992), Coarse Gaussian & Subspace Discriminant (AUC = 0.9986), Coarse KNN (AUC = 0.9961), Weighted KNN (AUC = 09940), Fine Gaussian (AUC = 0.9922), Quadratic SVM (AUC = 0.9887), Fine Gaussian (AUC = 0.9750) etc.

The results using different distance weight metrics denote that performance is different for each three cases as we increased the Neighbors. The performance though monotonically decreases; however, the inverse squared distance weight gives higher accuracy for Neighbors 8 to 100 as depicted in the Fig. 4 followed by the Inverse distance weight and equal distance weight. This gives us a new direction, that might be very helpful to choose the distance weight and Neighbors to get the improved performance instead of only using the default parameters or options used for any classifier.

Conclusion

Epilepsy is most neurological disorder in which millions of deaths occur each year. These signals are highly complex, multi-varying nature, nonstationary thus requires multidomain feature extracting strategies and most robust machine learning classifiers for their diagnostic and prognostic analysis.

Based on the nature of these signals, we proposed multidomain features based on varying characteristics of epileptic EEG signals ranging from time domain to frequency domain, statistical, complexity and wavelet based entropy measures. We employed the most robust machine learning classifiers such as Support Vector Machine Kernels, Decision Trees, K-nearest neighbors, Ensemble Classifiers with advanced parametrization approach to achieve higher accuracy and deeper learning for detecting the epileptic seizure. This automated approach can be used by clinicians for their diagnostic purpose to save millions of lives each year. We used the combined hybrid feature set for classification and detection purposes. The support Vector Machine classifiers gives more effective and optimal results at multiple kernel scales, parameterization, learning rates, while KNN also gives higher performance using different distance metrics and varying Neighbors. These methods with multidomain features gives outer performed results to distinguish the healthy and epileptic subjects than the traditional approaches. Moreover, the robust machine learning classifiers with multimodal features extracting approaches gives the highest evaluation performance to distinguish the postictal heart beat oscillations from epileptic seizure with ictal interval.

Limitations of Study and Future Recommendations

The research reported in this manuscript is focused on detecting the epileptic seizure during ictal and postictal intervals by extracting multimodal features and employing robust machine learning methods. The study revealed very interesting results, however, there are several limitations of this study as the number of subjects is small and lack of aging and gender base analysis. Furthermore, data set were taken from publicly available databases and clinical profile of data was not available. In future, we will employ these techniques and features extracting strategies to a larger database with respect to gender, age, seizure type and interictal intervals.

References

Acharya UR, Fujita H, Sudarshan VK, Bhat S, Koh JEW (2015) Application of entropies for automated diagnosis of epilepsy using EEG signals: a review. Knowl Based Syst 88:85–96

Adeli H, Zhou Z, Dadmehr N (2003) Analysis of EEG records in an epileptic patient using wavelet transform. J Neurosci Methods 123:69–87

Al-Aweel IC, Krishnamurthy KB, Hausdorff JM, Mietus JE, Ives JR, Blum AS, Schomer DLGAL (1999) Post-ictal heart rate oscillations in partial epilepsy. Neurology 53:1590–1592

Andrzejak RG, Lehnertz K, Mormann F, Rieke C, David P, Elger CE (2001) 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 Nonlinear Soft Matter Phys 64:61907

Avci E, Hanbay D, Varol A (2007) An expert discrete wavelet adaptive network based fuzzy inference system for digital modulation recognition. Expert Syst Appl 33:582–589

Ayoubian L, Lacoma H, Gotman J (2013) Automatic seizure detection in SEEG using high frequency activities in wavelet domain. Med Eng Phys 35:319–328

Bashar SK, Hassan AR, Bhuiyan MIH (2015) Identification of motor imagery movements from EEG signals using Dual Tree Complex Wavelet Transform. Int Conf Adv Comput Commun Inform 2015:290–296

Bashar SK, Hassan AR, Bhuiyan MIH (2016) Motor imagery movements classification using multivariate EMD and short time Fourier transform. In: 12th IEEE International conference on electronics, energy, environment, communication, computers, control (E3-C3), INDICON 2015, pp 1–6

Bengio Y, Lecun Y (2007) Scaling learning algorithms towards AI. Large Scale Kernel Mach 34:321–360

Bigger JT, Kleiger RE, Fleiss JL, Rolnitzky LM, Steinman RC, Miller JP (2016) Components of heart rate variability measured during healing of acute myocardial infarction. Am J Cardiol 9149:208–215

Bilchick KC, Fetics B, Djoukeng R, Gross Fisher S, Fletcher RD, Singh SN, Nevo E, Berger RD (2002) Prognostic value of heart rate variability in chronic congestive heart failure (Veterans Affairs' Survival Trial of Antiarrhythmic Therapy in Congestive Heart Failure). Am J Cardiol 90:24–28

Burges CJC (1998) A tutorial on support vector machines for pattern recognition. Data Min Knowl Discov 2:121–167

Casolo GC, Stroder P, Signorini C, Calzolari F, Zucchini M, Balli E, Sulla A, Lazzerini S (1992) Heart rate variability during the acute phase of myocardial infarction. Circulation 85:2073–2079

Choi A, Shin H (2017) Photoplethysmography sampling frequency: pilot assessment of how low can we go to analyze pulse rate variability with reliability? Physiol Meas 38:586–600

Chou KC (2005) Using amphiphilic pseudo amino acid composition to predict enzyme subfamily classes. Bioinformatics 21:10–19

Chou KC, Bin SH (2007) Recent progress in protein subcellular location prediction. Anal Biochem 370:1–16

Chou KC, Shen HB (2007a) Signal-CF: a subsite-coupled and window-fusing approach for predicting signal peptides. Biochem Biophys Res Commun 357:633–640

Chou K-C, Shen H-B (2007b) Euk-mPLoc: a fusion classifier for large-scale eukaryotic protein subcellular location prediction by incorporating multiple sites. J Proteome Res 6:1728–1734

Costa M, Goldberger AL, Peng C-KC-K (2002) Multiscale entropy analysis of complex physiologic time series. Phys Rev Lett 89:68102

Dheeba J, Albert Singh N, Tamil Selvi S (2014) Computer-aided detection of breast cancer on mammograms: a swarm intelligence optimized wavelet neural network approach. J Biomed Inform 49:45–52

Dobrowolski AP, Wierzbowski M, Tomczykiewicz K (2012) Multiresolution MUAPs decomposition and SVM-based analysis in the classification of neuromuscular disorders. Comput Methods Programs Biomed 107:393–403

Dodds KL, Miller CB, Kyle SD, Marshall NS, Gordon CJ (2017) Heart rate variability in insomnia patients: a critical review of the literature. Sleep Med Rev 33:88–100

Dragotti PL, Vetterli M (2003) Wavelet Footprints: theory. Algorithms Appl 51:1306–1323

Erfani SM, Baktashmotlagh M, Rajasegarar S, Karunasekera S, Leckie C (2015) R1SVM: a randomised nonlinear approach to large-scale anomaly detection. Proc Natl Conf Artif Intell 1:432–438

Esco MR, Williford HN, Flatt AA, Freeborn TJ, Nakamura FY (2017) Ultra-shortened time-domain HRV parameters at rest and following exercise in athletes: an alternative to frequency computation of sympathovagal balance. Eur J Appl Physiol 0:1–10

Faust O, Acharya UR, Adeli H, Adeli A (2015) Wavelet-based EEG processing for computer-aided seizure detection and epilepsy diagnosis. Seizure 26:56–64

Ferland RJ, Smith J, Papandrea D, Gracias J, Hains L, Kadiyala SB, O'Brien B, Kang EY, Beyer BS, Herron BJ (2017) Multidimensional genetic analysis of repeated seizures in the hybrid mouse diversity panel reveals a novel epileptogenesis susceptibility locus. G3 Genes Genomes Genet. https://doi.org/10.1534/g3.117.042234

Fu K, Qu J, Chai Y, Zou T (2015) Hilbert marginal spectrum analysis for automatic seizure detection in EEG signals. Biomed Signal Process Control 18:179–185

Fujita H, Acharya UR, Sudarshan VK, Ghista DN, Sree SV, Eugene LWJ, Koh JEW (2016) Sudden cardiac death (SCD) prediction based on nonlinear heart rate variability features and SCD index. Appl Soft Comput J 43:510–519

Gabor D (1946) Theory of communication. J Inst Electr Eng Part III Radio Commun Eng 93(26):429–457

Gajic D, Djurovic Z, Gligorijevic J, Di Gennaro S, Savic-Gajic I (2015) Detection of epileptiform activity in EEG signals based on time-frequency and non-linear analysis. Front Comput Neurosci 9:1–16

Gammerman A, Luo Z, Vega J, Vovk V (2016) Conformal and probabilistic prediction with applications: 5th international symposium, COPA 2016 Madrid, Spain, April 20–22, 2016 proceedings. Lect Notes Computer Science (including Subser Lect Notes Artif Intell Lect Notes Bioinformatics), vol 9653, pp 185–195

Geronikolou SA, Albanopoulos K, Chrousos G, Cokkinos D (2017) GeNeDis 2016. 987:249–259

Ghosh-Dastidar S, Adeli H, Dadmehr N (2007) Mixed-band wavelet-chaos-neural network methodology for epilepsy and epileptic seizure detection. IEEE Trans Biomed Eng 54:1545–1551

Glass L, Kaplan D (1993) Time series analysis of complex dynamics in physiology and medicine. Med Prog Technol 19:115

Goldberger AL, Amaral LAN, Glass L, Hausdorff JM, Ivanov PC, Mark RG, Mietus JE, Moody GB, Peng C-K, Stanley HE (2000) Components of a new research resource for complex physiologic signals. Circulation 101:e215–e220

Gotman J (1982) Automatic recognition of epileptic seizures in the EEG. Electroencephalogr Clin Neurophysiol 54:530–540

Gotman J, Gloor P (1976) Automatic recognition and quantification of interictal epileptic activity in the human scalp EEG. Electroencephalogr Clin Neurophysiol 41:513–529

Guo L, Rivero D, Seoane JA, Pazos A (2009) Classification of EEG signals using relative wavelet energy and artificial neural networks. In: Proceedings of ACM/SIGEVO Summit on Genetic and Evolutionary Computation—GEC'09, p 177

Guo L, Rivero D, Pazos A (2010) Epileptic seizure detection using multiwavelet transform based approximate entropy and artificial neural networks. J Neurosci Methods 193:156–163

Guo L, Rivero D, Dorado J, Munteanu CR, Pazos A (2011) Automatic feature extraction using genetic programming: an application to epileptic EEG classification. Expert Syst Appl 38:10425–10436

Hajian-Tilaki K (2013) Receiver operating characteristic (ROC) curve analysis for medical diagnostic test evaluation. Casp J Intern Med 4:627–635

Hassan AR (2015) Automatic screening of Obstructive Sleep Apnea from single-lead Electrocardiogram. In: 2nd International conference on electrical engineering and information communication technology (iCEEiCT 2015), pp 21–23

Hassan AR (2016) A comparative study of various classifiers for automated sleep apnea screening based on single-lead electrocardiogram. In: ICEEE 2015—1st international conference on electrical and electronic engineering, pp 45–48

Hassan AR, Bhuiyan MIH (2016) Dual tree complex wavelet transform for sleep state identification from single channel electroencephalogram. In: 1st IEEE international conference on telecommunications and photonics (ICTP 2015), pp 1–5

Hassan AR, Bhuiyan MIH (2017a) Automated identification of sleep states from EEG signals by means of ensemble empirical mode decomposition and random under sampling boosting. Comput Methods Programs Biomed 140:201–210

Hassan AR, Bhuiyan MIH (2017b) An automated method for sleep staging from EEG signals using normal inverse Gaussian parameters and adaptive boosting. Neurocomputing 219:76–87

Hassan AR, Haque MA (2016) Computer-aided obstructive sleep apnea screening from single-lead electrocardiogram using statistical and spectral features and bootstrap aggregating. Biocybern Biomed Eng 36:256–266

Hassan AR, Haque MA (2017) An expert system for automated identification of obstructive sleep apnea from single-lead ECG using random under sampling boosting. Neurocomputing 235:122–130

Hassan AR, Bhuiyan MIH (2016b) Automatic sleep scoring using statistical features in the EMD domain and ensemble methods. Biocybern Biomed Eng 36:248–255

Hassan AR, Subasi A (2016) Automatic identification of epileptic seizures from EEG signals using linear programming boosting. Comput Methods Programs Biomed 136:65–77

Hassan AR, Subasi A (2017) A decision support system for automated identification of sleep stages from single-channel EEG signals. Knowl Based Syst. https://doi.org/10.1016/j.knosys.2017.05.005

Hassan AR, Siuly S, Zhang Y (2016) Epileptic seizure detection in EEG signals using tunable-Q factor wavelet transform and bootstrap aggregating. Comput Methods Programs Biomed 137:247–259

Hayat M, Khan A (2012) Discriminating outer membrane proteins with fuzzy k-nearest neighbor algorithms based on the general form of Chou's PseAAC. Protein Pept Lett 19:411–421

Ho KKL, Moody GB, Peng C-K, Mietus JE, Larson MG, Levy D, Goldberger AL (1997) Predicting survival in heart failure case and control subjects by use of fully automated methods for deriving nonlinear and conventional indices of heart rate dynamics. Circulation 96:842–848

Huang FJ, LeCun Y (2006) Large-scale learning with SVM and convolutional nets for generic object categorization. Proc IEEE Comput Soc Conf Comput Vis Pattern Recognit 1:284–291

Hussain L, Aziz W (2016) Time-frequency spatial wavelet phase coherence analysis of EEG in EC and EO during Resting state. Proc Comput Sci 95:297–302

Hussain L, Aziz W, Kazmi SZH, Awan IA (2014) Classification of human faces and non faces using machine learning techniques. Int J Electron Electr Eng 2:116–123

Hussain L, Aziz W, Khan AS, Abbasi AQ, Hassan SZ (2015) Classification of electroencephlography (EEG) alcoholic and control subjects using machine learning ensemble methods. J Multidiscip Eng Sci Technol 2:126–131

Hussain L, Shafi I, Saeed S, Abbas A, Awan IA, Nadeem SA, Kazmi SZH, Shah SA (2017a) A radial base neural network approach for emotion recognition in human speech. Int J Comput Sci Netw, Secur, p 17

Hussain L, Aziz W, Saeed S, Shah SA, Nadeem MSA, Awan IA, Abbas A, Majid A, Kazmi SZH (2017b) Quantifying the dynamics of electroencephalographic (EEG) signals to distinguish alcoholic and non-alcoholic subjects using an MSE based K-d tree algorithm. Biomed Eng. https://doi.org/10.1515/bmt-2017-0041

Hussain L, Aziz W, Alowibdi JS, Habib N, Rafique M, Saeed S, Kazmi SZH (2017c) Symbolic time series analysis of electroencephalographic (EEG) epileptic seizure and brain dynamics with eye-open and eye-closed subjects during resting states. J Physiol Anthropol 36:21

Iscan Z, Dokur Z, Demiralp T (2011) Classification of electroencephalogram signals with combined time and frequency features. Expert Syst Appl 38:10499–10505

Jaiswal AK, Banka H (2017) Local pattern transformation based feature extraction techniques for classification of epileptic EEG signals. Biomed Signal Process Control 34:81–92

James CJ, Eng BE (1997) Detection of epileptiform activity in the electroencephalogram using artificial neural networks

Kaleem M, Guergachi A, Krishnan S (2013) EEG seizure detection and epilepsy diagnosis using a novel variation of Empirical Mode Decomposition. In: Proceedings of Annual International Conference of the IEEE Engineering in Medicine and Biology Society(EMBS), pp 4314–4317

Kang JH, Chung YG, Kim SP (2015) An efficient detection of epileptic seizure by differentiation and spectral analysis of electroencephalograms. Comput Biol Med 66:352–356

Kannathal N, Choo ML, Acharya UR, Sadasivan PK (2005) Entropies for detection of epilepsy in EEG. Comput Methods Programs Biomed 80:187–194

Kaya Y, Uyar M, Tekin R, Yıldırım S (2014) 1D-local binary pattern based feature extraction for classification of epileptic EEG signals. Appl Math Comput 243:209–219

Kleiger RE, Miller JP, Bigger JT, Moss AJ (1987) Decreased heart rate variability and its association with increased mortality after acute myocardial infarction. Am J Cardiol 59:258–282

Kuang D, Yang R, Chen X, Lao G, Wu F, Huang X, Lv R, Zhang L, Song C, Ou S (2017) Depression recognition according to heart rate variability using Bayesian Networks. J Psychiatr Res 95:282–287

Lee SH, Lim JS, Kim JK, Yang J, Lee Y (2014) Classification of normal and epileptic seizure EEG signals using wavelet transform, phase-space reconstruction, and Euclidean distance. Comput Methods Programs Biomed 116:10–25

Lewis GJ, Bates TC (2013) The long reach of the gene. Psychologist 26:194–198

Li Y, Wee CY, Jie B, Peng Z, Shen D (2014) Sparse multivariate autoregressive modeling for mild cognitive impairment classification. Neuroinformatics 12:455–469

Malik M (1996) Task force of the European society of cardiology and the north American society of pacing and electrophysiology. Heart rate variability. Standards of measurement, physiological interpretation, and clinical use. Eur Heart J 17:354–381

Mormann F, Andrzejak RG, Elger CE, Lehnertz K (2007) Seizure prediction: the long and winding road. Brain 130:314–333

Muller KR, Mika S, Ratsch G, Tsuda K, Scholkopf B (2001) An introduction to kernel-based learning algorithms. IEEE Trans Neural Netw 12:181–201

Nigam VP, Graupe D (2004) A neural-network-based detection of epilepsy. Neurol Res 26:55–60

Niknazar M, Mousavi SR, Vosoughi Vahdat B, Sayyah M (2013) A new framework based on recurrence quantification analysis for epileptic seizure detection. IEEE J Biomed Heal Informatics 17:572–578

Ocak H (2009) Automatic detection of epileptic seizures in EEG using discrete wavelet transform and approximate entropy. Expert Syst Appl 36:2027–2036

Orhan U, Hekim M, Ozer M (2011) EEG signals classification using the K-means clustering and a multilayer perceptron neural network model. Expert Syst Appl 38:13475–13481

Pan YH, Lin WY, Wang YH, Lee KT (2011) Computing multiscale entropy with orthogonal range search. J Mar Sci Technol 19:107–113

Peker M, Sen B, Delen D (2016) A novel method for automated diagnosis of epilepsy using complex-valued classifiers. IEEE J Biomed Heal Informatics 20:108–118

Pincus SM (1991) Approximate entropy as a measure of system complexity. Mathematics 88:2297–2301

Polat K, Güneş S (2007) Classification of epileptiform EEG using a hybrid system based on decision tree classifier and fast Fourier transform. Appl Math Comput 187:1017–1026

Ponikowski P, Anker SD, Chua TP et al (1997) Depressed heart rate variability as an independent predictor of death in chronic congestive heart failure secondary to ischemic or idiopathic dilated cardiomyopathy. Am J Cardiol 79:1645–1650

Rajendra Acharya U, Vinitha Sree S, Alvin APC, Suri JS (2012) Use of principal component analysis for automatic classification of epileptic EEG activities in wavelet framework. Expert Syst Appl 39:9072–9078

Rathore S, Hussain M, Aksam Iftikhar M, Jalil A (2014) Ensemble classification of colon biopsy images based on information rich hybrid features. Comput Biol Med 47:76–92

Rich MW, Saini JS, Kleiger RE, Carney RM, Freedland KE (1988) Correlation of heart rate variability with clinical and angiographic variables and late mortality after coronary angiography. Am J Cardiol 62:714–717

Rivero D, Fernandez-Blanco E, Dorado J, Pazos A (2011) A new signal classification technique by means of Genetic Algorithms and kNN. In: IEEE Congress on Evolutionary Computation (CEC), pp 581–586

Rosso OA, Blanco S, Yordanova J, Kolev V, Figliola A, Schürmann M, Başar E (2001) Wavelet entropy: a new tool for analysis of short duration brain electrical signals. J Neurosci Methods 105:65–75

Rouslan A (2008) www.econstor.eu

Samiee K, Kovács P, Gabbouj M (2015) Epileptic seizure classification of EEG time-series using rational discrete short-time fourier transform. IEEE Trans Biomed Eng 62:541–552

Schwenker F, Trentin E (2014) Pattern classification and clustering: a review of partially supervised learning approaches. Pattern Recognit Lett 37:4–14

Seely AJE, Macklem PT (2004) Complex systems and the technology of variability analysis. Crit Care 8:R367–R384

Sima CA, Inskip JA, Sheel AW, van Eeden SF, Reid WD, Camp PG (2017) The reliability of short-term measurement of heart rate variability during spontaneous breathing in people with chronic obstructive pulmonary disease. Rev Port Pneumol 23:338–342

Srinivasan V, Eswaran C, Sriraam N (2007) Approximate entropy-based epileptic EEG detection using artificial neural networks. IEEE Trans Inf Technol Biomed 11:288–295

Stanley Raj A, Hudson Oliver D, Srinivas Y, Viswanath J (2016) Wavelet denoising algorithm to refine noisy geoelectrical data for versatile inversion. Model Earth Syst Environ 2:36

Stochholm A, Mikkelsen K, Kidmose P (2016) Automatic sleep stage classification using ear-EEG. In: Proceedings of the 20th annual international conference of the IEEE engineering in medicine and biology society (EMBS 2016) pp 4751–4754

Subasi A (2007) EEG signal classification using wavelet feature extraction and a mixture of expert model. Expert Syst Appl 32:1084–1093

Subasi A (2013) Classification of EMG signals using PSO optimized SVM for diagnosis of neuromuscular disorders. Comput Biol Med 43:576–586

Subasi A, Erçelebi E (2005) Classification of EEG signals using neural network and logistic regression. Comput Methods Programs Biomed 78:87–99

Subasi A, Kevric J, Abdullah Canbaz M (2017) Epileptic seizure detection using hybrid machine learning methods. Neural Comput Appl 28:1–9

Takeda M, Ina H, Kobayashi S (1982) Fourier-transform method of fringe-pattern analysis for computer-based topography and interferometry. J Opt Soc Am 72:156

Tuininga YS, van Veldhuisen DJ, Brouwer J, Haaksma J, Crijns HJ, Man in't Veld AJ, Lie KI (1994) Heart rate variability in left ventricular dysfunction and heart failure: effects and implications of drug treatment. Br Heart J 72:509–513

Tzallas AT, Tsipouras MG, Fotiadis DI (2007) Automatic seizure detection based on time-frequency analysis and artificial neural networks. Comput Intell Neurosci. https://doi.org/10.1155/2007/80510

Tzallas AT, Tsipouras MG, Tsalikakis DG, Karvounis EC, Astrakas L, Konitsiotis S, Tzaphlidou M (2009) Automated epileptic seizure detection methods: a review study (Chap 4). Epilepsy Histol Electroencephalogr Psychol Asp 75–98

Tzallas AT, Tsipouras MG, Fotiadis DI, Member S (2009b) Epileptic seizure detection in EEGS using time—frequency. Analysis 13:703–710

Übeyli ED (2010) Least squares support vector machine employing model-based methods coefficients for analysis of EEG signals. Expert Syst Appl 37:233–239

Van Hoogenhuyze D, Weinstein N, Martin GJ, Weiss JS, Schaad JW, Sahyouni XN, Fintel D, Remme WJ, Singer DH (1991) Reproducibility and relation to mean heart rate of heart rate variability in normal subjects and in patients with congestive heart failure secondary to coronary artery disease. Am J Cardiol 68:1668–1676

Vapnik VN (1999) An overview of statistical learning theory. IEEE Trans Neural Netw 10:988–999

Vempati S, Vedaldi A, Zisserman A, Jawahar CV (2010) Generalized RBF feature maps for efficient detection. In: Procedings of British machine vision conference, pp 2.1–2.11

Walters-Williams J, Li Y (2011) A New Approach to Denoising EEG Signals-Merger of Translation Invariant Wavelet and ICA. Int J Biom Bioinform 5:130–148

Wang D, Miao D, Xie C (2011) Best basis-based wavelet packet entropy feature extraction and hierarchical EEG classification for epileptic detection. Expert Syst Appl 38:14314–14320

Wang R, Kwong S, Wang X, Jiang Q (2015) Continuous valued attributes. IEEE Trans Cybern 45:1262–1275

Wang L, Xue W, Li Y, Luo M, Huang J, Cui W, Huang C (2017) Automatic epileptic seizure detection in EEG signals using multi-domain feature extraction and nonlinear analysis. Entropy 19:1–17

Wu Y, Zhou Y, Saveriades G, Agaian S, Noonan JP, Natarajan P (2013) Local Shannon entropy measure with statistical tests for image randomness. Inf Sci (Ny) 222:323–342

Zhang P, Gao BJ, Zhu X, Guo L (2011) Enabling fast lazy learning for data streams. In: Proceedings of IEEE international conference on data mining (ICDM), pp 932–941