How to Cite:
Meshram, P. S., & Gharpure, D. C. (2022). Detection of epileptic seizures on EEG signals using Decision Tree, KNN, SVM and ensemble classifiers. *International Journal of Health Sciences*, *6*(S2), 13528–13556. https://doi.org/10.53730/ijhs.v6nS2.8623

Detection of epileptic seizures on EEG signals using Decision Tree, KNN, SVM and ensemble classifiers

Pallavi S. Meshram
Assistant Professor, Department of Electronic & Instrumentation Science, Savitribai Phule Pune University, Pune, Maharashtra, India
Corresponding author email: pallavimeshram9@gmail.com

Damayanti C. Gharpure
Professor, Department of Electronic & Instrumentation Science, Savitribai Phule Pune University, Pune, Maharashtra, India
Email: dcgunipune@gmail.com

***Abstract*---**Epilepsy is a neurological condition resulting to brain cell stimulation. According to the findings of the research, an electroencephalogram (EEG) can identify epileptic episodes in patients who have epilepsy. Performance evaluations based on EEG detection of epilepsy require feature extraction methods. As a result of our research, we were able to identify a number of different feature extraction approaches. These methodologies were dependent upon nonlinear, wavelet-based entropy characteristics, time - frequency domain characteristics, and a few statistical traits. An additional in-depth examination that made use of cutting-edge machine learning classifiers as well as numerous factors was carried out. When evaluating kernels for support vector machines, the multiclass kernel as well as the box constraint levels are both useful tools. In a similar manner, we computed the various distance measures, neighbour weights, and neighbour relationships for the K-nearest neighbours (KNN). In a similar manner, we altered the decision trees of the paramours depending on maximal splits as well as split criteria, and we evaluated the ensemble classifiers based on a variety of ensemble approaches and learning rates. TPR, NPR, PPV, accuracy, & AUC were the metrics that were utilised to evaluate performance, and tenfold validation data was being used for both training and testing. Through this investigation, do an in analytical method was used to extract any number of factors in order to make more use powerful machine learning classifiers though merged with more complex optimal alternatives. This was done so as to better understand the connection among the elements. The Support Vector Machine linear kernel as well as the KNN classifier with both the City block distance measure produced a high accuracy rate of 99.5 percent, and was higher than with the accuracy achieved and using the classifiers' default parameters. This indicates that the Support Vector Machine linear kernel as well as the KNN classifier with City block distance metric are superior. This was due to the fact that both of these classifiers took into account the distance between adjacent city blocks. In addition, applying SVM with varied kernel sizes allows for the determination of the best separation (AUC = 0.9991, 0.9990). In addition to this, those K-nearest neighbours with just an inverse squared distance weight perform much better at different Neighbors. In addition, a maximum performance of one hundred percent was achieved by employing a variety of distinct machine learning classifiers to differentiate between epileptic ictal patients and the postictal heart rate oscillations of postictal patients. This was done in order to determine which patients were having an epileptic ictal episode.

***Keywords*---**epilepsy, support vector machine kernels, electroencephalogram, ensemble classifier, machine learning.

Introduction

One of the most prevalent neurological conditions, epilepsy affects between 1 & 2 two percent of world's population. (Mormann et al. 2007). Sudden accidental death in epilepsy (SUDEP) occurs when epileptic patients with or without evidence of seizure die suddenly, unwitnessed, nondrowning, and nontraumatically. Epilepsy is a neurological issue that impacts approximately 1.5 billion people. After stroke, this is the neurological condition that is seen most frequently.

Epileptic seizures are brought on by a sudden malfunction as well as synchronisation of this group of neurons. As a consequence of this, excessive & hypersynchronous neuronal activity occurs in the brain, which is not only the primary reason for epileptic seizures but also the root cause of a disorder. Epilepsy is characterised by recurrent seizures, which are often referred to as epileptic seizures. Recurrent seizures are among the disease's hallmarks and are regarded to be among the disease's distinguishing characteristics.The clinical presentation of these seizures can indicate that they are either extensive, focal, partial, unclassified, or unilateral. Other possible manifestations include unclassified seizures. Seizures that are focused impact just one hemisphere of brain, and also the symptoms of such seizures express themselves in parts of the human body and mental processes which correspond to the hemisphere that is being affected. In addition, when a person experiences a generalised epileptic seizure, one‘s entire body has been affected, that causes them to lose consciousness and also have motor symptoms that affect all sides of their body. During a generalised epileptic seizure, a person may also experience loss of sensation in their extremities. These fits can occur in people of any age, but they are more prevalent in people who are either younger or older (Hassan et al. 2016). The researchers are carrying on with their efforts to improve their capacity to provides health epileptic episodes as well as accurately anticipate when they will occur.

Visual scanning of electroencephalogram (EEG) records has traditionally been used to identify epileptic convulsions and spikes; however, this method is time-consuming, expensive, and prone to error (Ocak, 2009). It is required to design a system which is both efficient and effective for identifying epileptic activity in EEG data so as to discover a solution to this problem. Doing so will allow for the possibility of resolving the issue (Hassan et al., 2016) (Tzallas et al., 2007); (Guo et al., 2011) (Fu et al., 2015). In an effort to keep track of any telltale indicators that the brain may be undergoing epileptic processes, researchers have made use of a wide variety of different automated methods. In addition, observers have utilized computerised detection methods (Gotman, 1982) (Gloor, 1976) to identify epileptic events that are able to be located in the EEG. This was done in order to determine whether or not epileptic episodes may be found in the EEG. By utilising the wavelet technique, (Adeli et al., 2003) were able to conduct a kinetics analysis of the epileptic signals found in EEG data. (Ican et al., 2011)(2011) successfully recovered the EEG epileptic signals by using time and frequency components in their analysis. (Guo et al., 2010) categorised epileptic signals by utilising the multiwavelet transform in conjunction with approximation entropy as well as artificial neural networks. Methods such as time–frequency analysis, wavelet analysis, entropy analysis, and feature extraction using artificial neural networks (ANN) have also been utilized (Graupe, 2004) (Orhan et al., 2011) classified indications of epileptic activity in the EEG obtained through the application of artificial neural networks in conjunction using “K-means clustering”. Wavelet transform & “artificial neural networks”, as detailed by (Kang et al., 2015) (Ghosh-Dastidar, 2007), can be used to identify EEG recordings of epileptic episodes.

(Bashar et al., 2016) utilised multivariate EMD in addition to the limited period Fourier transform with their most recent piece of research. Both of these approaches were successful in accomplishing what they set out to do. On the other hand, (Bashar et al., 2015)distinguished among EEG left-hand movement & right-hand movement by employing a dual complex wavelet transform. This was done in order to analyse the data. Additionally, the extraction of statistical & spectral data was put to use in order to carry out automatic detection and classification of naps. This was done in order to better understand naps. (Stochholm et al., 2016) utilised the use of adaptive improving classifiers, whilst (Hassan et al., 2016) acquired statistical parameters from an endless amount of intrinsic mode function (IMF) utilising a huge number of machine learning classifiers. Extraction of statistical & spectral features from EMD has been done, in fact, by (Haque, 2016) using Bootstrapping aggregate & ensembles approaches, respectively. (Bhuiyan, 2016,2017) also used Bootstrapping aggregate & ensembles approaches. They also made use of such a inverted Gaussian variable in addition to adaptable boosting in a manner that was equivalent. In a similar manner, (Subasi, 2017) employed the tuneable Q wavelet (often abbreviated as TQWT).(Fig. 1).

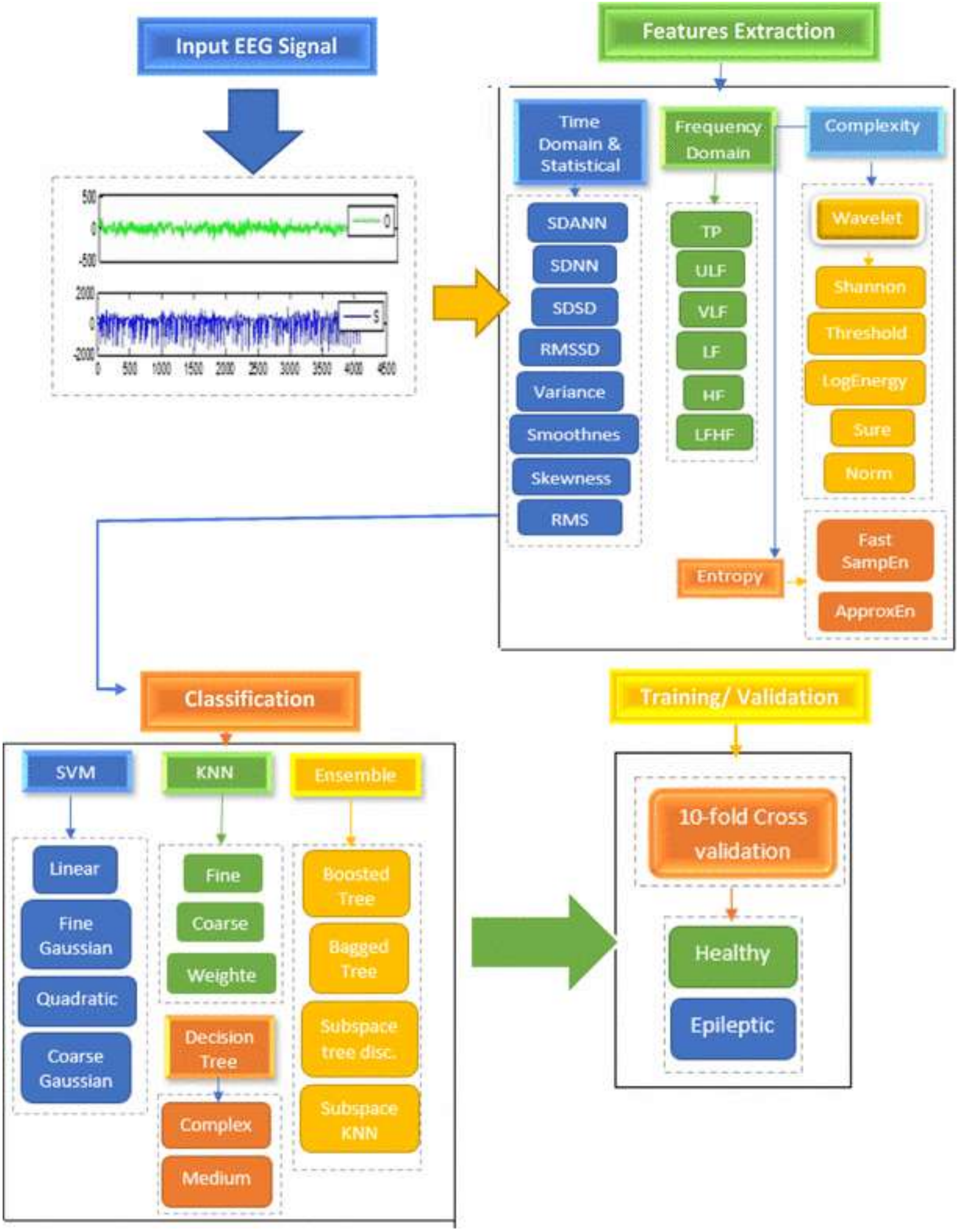

Figure 1A schematic representation of the epileptic seizure detection process, which includes the collection of several feature extracting methodologies and multi-criteria machine learning classifiers

Using several feature extraction methods such as nonlinear dynamical measures, multi-domain features with quick entropy measurement via KD tree algorithms method, and wavelet entropy to boost detection performance, we employed automatic systems to determine epileptic activity in EEG data. First reliable machine learning techniques were utilised to fine-tune the variables that surpassed the prior approaches. For the purpose of variability analysis, time domain features are obtained. These are the identical qualities that are commonly utilised in a broad range of applications, including brain and the heart, amongst others. In order to obtain a handle on the frequency component, it is necessary to first collect the characteristics of a frequency domain and then utilise those characteristics in a diverse variety of apps of spectrum analysis. In order to extract the attributes of nonlinear dynamics while maintaining a high level of performance, a KD tree approach technique is utilised. This method would be much more efficient in terms of speed and memory performance, and that is what is used here. These characteristics include the approximation entropy as well as the sample entropy, which are both known as entropies. Wavelet entropy can be broken down into its component parts, which include Shannon entropy, threshold, log energy, & sure. In addition, one of the features is absolute certainty. Other statistical qualities, like smoothness, skewness, and kurtosis, as well as root mean square to mean, can be obtained through the use of the same method. Modern machine learning classifiers require different feature combinations as input in order to correctly categorise the data they are given. We used multiscale kernels and box constraint levels while selecting support vector classification kernels in order to increase assessment performance compared to when we used the default settings. This was done in place of the default settings. We utilised a wide range of distance computation measures when working on the K-nearest neighbours approach. In a similar manner, the effectiveness of KNN was evaluated by employing a variety of Neighbors & distance weights, including equal, inverse, & squared inverse weights, and others as well. This was done in order to determine the optimal combination of these two factors. Changing these parameters has an impact on the performance assessment taken as a whole. The fundamental parameters of decision trees were decided based on a wide range of considerations, such as the maximum number of splits, the split criterion, as well as a great many more. In a manner analogous to this, the performance was judged to use a variety of ensemble methods, each of which had a unique number of learners and unique learning rates for those learners.

Materials and methodology

Researchers were able to access the data set since the Department of Epileptology at Bonn University made it available to them (https://www.ukbonn.de/epileptologie/). The data set was collected from a database that is available to the general public. In their analysis of the data set, Andrzejak et al. provide a comprehensive and in-depth explanation (2001). Throughout the entirety of EEG data collection process, a setup consisting of 128 channel amplifiers and an average size reference are utilised. Throughout the entirety of the process of converting the information into a digital representation, a sampling frequency of 173.61 Hz and just a resolution of 12 bits are both utilised. During the course of this process, the data was translated into a digital format. The apparatus that was used to collect the data had such a spectral bandwidth which varied from 0.5 to 85 Hz, and so this bandwidth was the one that was utilised to collect the data. The complete collection of information gleaned from the EEG. Each set contains one hundred EEG segments, with Z, O, N, F, and S constituting those segments, and the duration of every segment was precisely 23.6 seconds. Sets O & Z was collected in five healthy volunteers while they were awake, with eye closed (for set O) and with another eye open (for set Z). The technique of electrode insertion that was used was one has been consistent throughout the entire study. Set O was collected with one eye closed, and set Z was collected with one eye open. The purpose of this was to collect data for both groups, hence it was doneSets N, F, & S were extracted from an EEG data archive that had been completed prior to the procedure. This archive was combed through in order to extract the sets. Set F segments was taken from epileptogenic zone of the brain, whereas set N segments came from the development of the hippocampus in the cerebral hemisphere that was healthy and unaffected by the disease. Both set of segments were taken from the same person in order to ensure continuity. On the other hand, sets N and F contained EEG recordings which were taken at times when seizure activity wasn't really present, and set S consisted of EEG recordings which were recorded during times when seizure action was present. We make use of electroencephalogram (EEG) data collected from epileptic sufferers in addition to healthy participants (Set O) so that we might carry out future research using this information. (Set S).

A further set of data on postictal pulse rate oscillation of partial epilepsy was extracted from a database that is open to the general public. This database includes 11 partial seizures which were recorded for 5 female patients who have been continuously monitored with EEG, ECG, and video. These patients were all under continuous observation. These patients were subjected to ongoing EEG and ECG monitoring, as well as video surveillance. Patients' ages ranged from 31 - 48, but they all suffered from partial seizures, the causes of which were localised in either the frontal or the temporal lobes. Although not all of the patients had the symptoms, several of them did exhibit signs of future generalisation. The recordings were produced in a way that was in accordance with a technique that had been devised by Israel Deaconess Medical Center. This indicates that its recordings were created in a manner that was not only trustworthy but also accurate. An offline data analysis were successfully finished with the help of some specialised tools. They were able to visually recognise the start and offset of seizures to an accuracy of within 0.1 seconds, all while a experienced electroencephalographer (D.L.S.) was merely being masked during heart rate variability study. The signals from the electrocardiogram were sampled at a rate of two hundred hertz (Hz) so that a continuous signal could be formed.With the use of arrhythmia analysis software, annotations of the each heart beat were produced. This programme is easily accessible to the general public. Changes in higher order non-stationary heart rate were excluded from the analysis due to the risk that these changes could obscure oscillations occurring at lower frequency. A least square filter of a fourth degree polynomial was utilised in order to detrend the time series. This is done in order to achieve the desired results. During the calculation, a rapid Fourier transform was performed on a rectangle frame, which allowed for the determination of spectral density.

Because they describe a pattern of change that what a live system goes through, biological signals are essential for understanding not only the dynamic behavior of healthy biological systems but how pathological disruption and ageing affects this same resilience of biological systems. This is because biological signals are what decide the general pattern that a network of living beings goes through. As a result of developments of cutting-edge technology including such computerised data collection systems, sophisticated monitoring devices, & instrumentational technologies, the field of biomedical signal processing is gaining more and more popularity. In the last two decades, there has been a discernible trend among biomedical researchers towards an increased interest in the development of novel strategies and instruments for the processing of biological signals. The physiological rhythmic changes linked to the condition have been identified by physicians. The human eye is an amazing pattern recognition system that can analyse extremely complex ECG and EEG data, and clinicians utilise it to diagnose a variety of medical conditions. On either hand, a comprehensive investigation of variability can serve as a gauge of the integrity of the underlying systems which are accountable for the generation of the dynamics. The spatial or temporal structure of the this complex system—changes in patterns of interconnection throughout terms of coverage & variance through sequence over time and the amount of variability—represent highly essential methodologies that can be used to treat and prognosticate patients. These methodologies can be found in the following sentence. One way to conceptualise these alterations is as a form of connectedness. Two distinct classes can be distinguished amongst these architectural elements: connectedness and variability. It has been determined that a benign epileptic episode includes ictal intervals & postictal heart rate oscillations can produce multimodal characteristics. These characteristics are based on variable, nonlinear, & nonstationary dynamics.

Quantifying the variability in HRV requires manipulating a number of factors that influence HRV indices in healthy people. The head being tilted back, standing (which boosts sympathetic activity), and taking slow, deep breaths are all examples of these characteristics (respiratory rate induced HRV is increased). In addition, the diversity of brain signals, and more specifically the variances with in temporal transit of brain signals, provides highly essential information regarding overall dynamics of the network. Because of the influence of variability, transitions between distinct functional network configurations can take place more easily in complex dynamical systems like the brain, regardless of whether an external stimulus is present or not.

In a similar vein, physiological signals can be conceptualised as a series of sinusoidal oscillations that occur at varying frequencies. In order to move from time domain to frequency domain domain, they make use of Fourier transform, which was first devised by a mathematician in the year 1807. This transform was named after Joseph Fourier. The process of changing data from time domain to frequency domain domain often makes use of two methods: the wavelet transform and the Hilbert transform. These are the two approaches that are utilised most commonly in the processBy using this method, the amplitude of every cosine and sine wave may be calculated on the basis of frequency of wave. This technique is known as spectrum analysis, and its name comes from the fact that it determines the magnitude (amplitude) of each frequency that is contributing to the overall signal. Waves with a cosine component and sine component are two examples of periodic waves.

In time series analysis, the standard deviation (SDNN), the standard deviation of an average interval (SDANN), as well as the square root of mean squared deviations over successive NN intervals (RMSSD) have been used to investigate global variation, long-term variation, & short-term variation, respectively. These three measures were each used in conjunction with the standard deviation of successive NN intervals. These values were determined by calculating the mean squared deviations across all of the NN intervals that came before them. Such time domain metrics have found widespread use in the diagnosis of such an increased risk of death in patients who are suffering from heart issues, dilated cardiomyopathy, & individuals who have just undergone myocardial infarction. In addition, in order to recognise epileptic episodes, a variety of different automated methods, such as in the time domain, as well as Fourier spectrum analysis for the purpose of removing features, frequency domain methods, DWT-based methods, as well as fast-Fourier transforms are utilised. The reason for this is to ensure that the identification process is carried out with the utmost precision.

Signal preprocessing

In many cases, low & weak frequency signals are adversely affected by complex lower frequencies noise, including such system interference. Before looking for epileptic seizures and identifying them through analysis of EEG data, the data are first preprocessed to eliminate noise. This research makes use of the wavelet threshold denoising strategy, which is superior than that of the Fourier transform denoising technique in terms of its performance. Because of its good local approximation capability, the Daudechies (db4) wavelet from the fourth order was selected as the wavelet of choice. This wavelet functions more effectively when applied to nonstationary data. In addition to that, a main component evaluation had done so as to reduce the total dimensions and then get rid of any superfluous characteristics.

The wavelet threshold approach was used to denoise the nonstationary signals:

$$\lambda = \delta\sqrt{2\log N}$$

Where is the wavelet threshold, where is the standard deviation, where is N, where is the length of a sample signal. When it comes to reducing white noise from just an EEG sample, the wavelet threshold method is by far the most reliable method.

Feature extraction

The epileptic seizure was detected using a multimodal feature extraction technique in this investigation. The properties of nonlinearity and nonstationarity are included in the extraordinarily complex nature of EEG signals. Qu et al. were able to effectively decompose EEG signals into a sequence of intrinsic mode functions (IMFs) thanks to the application using empirical mode decomposition.

These functions could be utilized to define the dynamics of nonstationary and nonlinear systems. Decomposing the EEG data utilizing empirical mode decomposition allowed for this to be accomplished (EMD). The characteristics of the IMF, which are derived from EMD, serve as the primary source of information for determining the nature and severity of epileptic seizures. The Support Vector Machine is employed for this purpose. The findings of the IMF indicate a high level of accuracy in categorization. However, some research suggests that the best way to classify epileptic episodes is to use multimodal features, which mix characteristics from multiple domains with nonlinear characteristics. Because of this, a consistent framework will be provided for combining the benefits that various EEG signal characteristics offer. In this body of work, humans also obtained characteristics for distinguishing epileptic seizure patients than in healthy individuals & postictal heart rate oscillations through utilising techniques in the time domain, the frequency domain, complexity-based measures, & wavelet entropy. This was done by analysing the data in the time domain, the frequency domain, and the wavelet entropy. This was done in order to acquire the characteristics. The implementation of these techniques has begun. Aside from this, during the course of our research, we were able to retrieve nonlinear features by employing sample entropy based solely on the KD tree algorithmic technique, as well as approximation entropy. In view of the fact that we've already been able to get this information, there was an additional benefit. In addition to this, we were successful in recovering these characteristics by employing an approach known as approximation entropy.

Linear methods

When it comes to capturing the temporal fluctuations & spectrum data through physiological signals (such as an ECG or EEG) that have been altered by a variety of diseases, the techniques which operate in the time domain and that those who operate in the frequency domain, respectively, are also the ones that are used the most frequently. These techniques are the ones that are used most frequently because they operate in their respective domains. Ayoubian et al. (2013) culled a variety of characteristics from temporal domain, the frequency domain, as well as the time–frequency domain in order to identify epileptic episodes. [Further citation is required] They were able to retrieve the HF activities as well as the maximum wavelet entropy, both are critical components of seizure start patterns with in HF activities. The seizure onset pattern in the HF activities were successfully retrieved. In addition, that information which is most helpful for identifying epileptic seizures could be discovered with in low frequency ranges, that stretch from 1 to 70 Hz. This information can be obtained with in low frequency ranges. When it comes to capturing the temporal fluctuations and spectrum information via physiological signals (such as an ECG or EEG) that have been altered by a variety of diseases, the techniques which operate in the time domain and those that operate in the frequency domain, respectively, are indeed the ones that are used the most frequently. These are the ones that are utilised most frequently because they are the ones that operate in their respective domains. Ayoubian et al. (2013) retrieved a number of features from temporal domain, the frequency domain, and the time–frequency domain in order to identify epileptic seizures. These characteristics were taken from the three different domains.

Time domain analysis

The examination of a segment in the time domain makes it possible to extract a number of different parameters, and the methods that make use of time domain are the simplest to put into practise. SDSD: determined using the standard deviation, the variance of the variations between successive time sequence periods in each segment. SDNN: variation of the future periods expressed as just a standard deviation for each section.

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

RMSSD: This is equal to the square root of the mean squared deviations for N consecutive EEG time - series data intervals.

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

SDANN: The standard deviation of a EEG intervals' averages.

$$SDANN = SD[u_1, u_2, u_3, u_4 \ldots u_n]$$

Frequency domain analysis

Time domain techniques, despite their simplicity, are unable to differentiate between parasympathetic and sympathetic contributions for HRV & EEG oscillations. The total power is calculated as the sum of all the EEG time-series data periods which have a frequency of up to 0.4 Hz (TP). In the field of electroencephalography (EEG), the phrase "very low frequency" corresponds to the sum spectral output for successive periods that fall between 0.003 and 0.04 Hz. This term came up because of the use of the term "very low frequency" in EEG (VLF).

The term "low frequency" refers to the total spectral intensity of all EEG waves which are consecutive and lie between 0.04 and 0.15 Hz. This is the definition of the term "low frequency." The total of all these spectral intensities is what is meant when people talk about "low frequency" (LF). It is referred to as high frequency, and is the word that is used to describe it, because the sum of the spectral energies for all of the EEG periods that are consecutive and occur between 0.15 and 0.4 Hz falls into this frequency range (HF).

The overall amount of spectral power exhibited by all the EEG's consecutive intervals, with such a frequency limit of 0.003 Hz. the ratio of power at low frequencies to power at high frequencies. the overall amount of spectral power exhibited by all of the EEG's consecutive intervals. This metric provides an analysis of the degree to which the sympathetic and parasympathetic neural systems were able to cooperate effectively inside the body as a whole.

Classification

The epileptic episode was detected and predicted using sophisticated machine learning methods. This model is used to estimate accuracy and other performance assessment metrics. The test sample's known label is compared with the model's categorised findings. Tenfold cross validation is utilised throughout the training, testing, & validation processes. The support vector machine (SVM) classifications is a great example of a method of supervised learning that is predicated on the theory of limited samples. The discipline of machine learning is where you can find this particular classifier. Traditional methods, which are based on the concept of empirical error minimization, were prone to overfitting in circumstances with a limited sample size. On the other hand, support vector machines, that are based on the concept of structure risk minimization, used to have a high potential for generalisation. The utility of proposed classification is demonstrated through the utilisation of a variety of classifiers, including Decision Trees, KNN, & Ensemble.

SVMs are more suited and can offer excellent generalisation even if the training set is biassed; since the loss function is convex, this results in a unique solution. The complexity of SVMs, which are nonparametric models, increases in a quadratic fashion as that of the number of records in the dataset grows. SVMs are even better for small datasets with a lot of feature. In order to achieve a high level of generalisation, an amount of training data must rise at a pace that is exponentially proportional to the number of variables. This is because the curse with dimensionality is caused when there are many more features than there are training samples. In addition, shallow architectures have certain practical limitations when it comes to properly modelling certain kinds of function families. In order to avoid these issues, models need to be developed that are capable of reflecting the broad scope of variation that may be present with in underlying data pattern without enumerating each of these variations individually. It is essential to utilise a condensed representation of data through order to grab a majority of variance through order to mitigate an issues caused by curse of dimensionality and to reduce the processing time of computing. This can be accomplished by using a compact representation of data.

Support Vector machine (SVM)

When it comes to classification utilising techniques of supervised learning, one of the most reliable algorithms that is now available is called the support vector machine (SVM). Support vector machines find widespread use across many different domains, such as pattern recognition, machine learning, & medical diagnostics, for name just a few of those domains (SVM). In addition, SVM is utilised in a broad variety of applications, some of which include diagnosis and characterisation, text recognition, content-based image processing, voice recognition, and biometrics, to name just a few of the many possible uses. Support vector machines, if applied to a space with just an infinite or large number of dimensions, provide a hyperplane and even collection of hyperplanes that could be used for categorization. [Case in point:] [Case in point:] [Case in point:] [Case in Because it has the most distance to go to reach the nearest training data point of any category, this hyperplane successfully achieves a fair separation (this distance is also known as functional margin). The bigger the margin, the less room there is for mistake in the generalisations that are created by the classifier when compared to when the margin was smaller. The objective of the Support Vector Machine is to pinpoint a hyperplane that offers a training example at the smallest feasible distance and in the shortest possible amount of time (SVM). When it comes to discussions on the SVM theory, this idea is typically referred to as the margin. It has come to light that hyperplane with the greatest potential for profits is also the one with the biggest margin. The support vector machine, often known as an SVM, has a variety of advantageous characteristics, one of the most notable of which being its greater generalisation performance in comparison to that of other methods. The support vector machine, often known as an SVM, is nothing more than a two-classifier that converts the input into the a hyperplane. This classifier works best with nonlinear training data or data with more dimensions.

The efficiency of an SVM classifier is contingent on a huge variety of factors coming together in the right combination. In order to identify the ideal parameter value, the grid search strategy was applied, and both grid range and also the step size were given a great deal of care. The constraint violation penalty which is associated with a data point which appears on just the wrong side of a decision surface has now been depicted by the linear kernel's single parameter, which would be the 'c' soft margin constant. This penalty is affiliated with a data point that appears just on the wrong side of the decision surface. This penalty is imposed on the data point whenever it appears on the opposite side of the decision surface from where it should be. A penalty of this kind is associated with a piece of data like that, which appears on the wrong side of the decision surface. The Support Vector Machine (SVM) that uses an RBF and Gaussian kernel function does, in fact, have two teaching parameters: cost (C), which controls the level of overfitting in the model; and sigma (), which regulates the level of nonlinearity in the model. Both of these teaching parameters can be found in the Support Vector Machine (SVM) manual. Both of these parameters can be found in the SVM's documentation. Both of these parameters can be adjusted during the training process. During training, either of these two parameters can have their values changed. The default cost function as well as sigma values were both used in this analysis.

Decision tree

The dataset is analysed using this Decision Tree classifier for recurring patterns of similarity, and the findings are then partitioned into a variety of categories. Wang et al. (2015) utilised decision trees in order to categorize data based on the selection of a characteristic which maximises & permanently fixes data division. This was accomplished by using the characteristic that was chosen. After that, the traits are dissected into the corresponding subcategories, and this process continues till all the prerequisites for termination have been satisfied. In order to develop a DT algorithm using mathematics, it is necessary to find solutions to the given set of equations.

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

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

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

where m is really the number of observations that can now be accessed, n is the number of variables that are considered to be independent, S is really the m-dimensional vector of variable that's been predicted from, I is the ith component of n-dimension independent variable, Xi is the pattern vector, & T is the transpose notation.

Ensemble classifiers

When classifying new cases utilising a variety of approaches, ensemble classifiers are constructed from a set of individual classifiers that have undergone individual training and then have their predictions combined. Learning algorithms that generate a group of classifiers and then use the aggregate weight of their predictions to categorise fresh data points are what this refers to. These methods have been successfully adopted to improve the accuracy of prediction in a variety of applications, includes, signal peptide prediction, protein, subcellular location prediction, and enzyme subfamily prediction, to name just a few of these applications. In a variety of settings, integrated classification algorithms typically exhibit superior performance over that of stand-alone classification methods. When the individual classifiers have been merged, the overall error rate has been reduced since an error made through one classifier might be compensated for an error made by another classifier; however, when the classifiers have been combined, the overall error rate has been reduced because of individual classifiers each get their own peculiarities and therefore can make a variety of errors during classification process.

K-nearest neighbour (KNN)

KNN is the method that is used the most commonly in a wide variety of fields, some of which include machine learning and pattern recognition, amongst others. KNN was utilised by Zhang et al. (2011) in order to overcome difficulties regarding classification. This technique is also known as the instance-based (lazy learning) algorithm in some circles. All of the training data samples have been stored, and they will be retrieved and when there are new observations which need to be categorised. This is the case even though a model or classifiers are not immediately being constructed. This aspect of the lazy learning method differentiates it from the eager learning strategy, which first creates classifiers before organising freshly observed data into categories. According to Schwenker & Trentin (2014), this strategy becomes even more significant in situations in which dynamic data needs to be adjusted and improved in a more expedient manner. KNN has made use of a number of different distance metrics.

The formulation of data for training and testing The Jack–Knife tenfold cross validation approach was utilised for the goals of creating training/testing data & optimising the parameters of the model. Using tenfold regression, which is one of the most well-known ways for validating the accuracy of a classifier, it is also one of the most widely used and effective methods. CV separates the data into tenfolds, of which ninefolds are used for training and the remaining three folds' sample classes are predicted using the knowledge gained from the training of the other ninefolds. The trained models have no information whatsoever about the test samples that are included within the test fold. Following the completion of the method 10 times, predictions are made for each class sample. At last, the accuracy of the classification is evaluated based on the anticipated labels of the samples that have not yet been seen. This operation is repeated for each possible configuration of the settings on the system, and the results of the categorization of the samples may be found in Tables 1, 2, 3, 4, 5, and 6.

Table 1 Evaluation of performance using matrices of confusion and comparison

	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%

Classifier	TNR (%)	TPR (%)	NPV (%)	PPV (%)	Overall accuracy (%)	AUC	95% CI	
							LB	UB
Support vector machine (SVM)								
Linear	98	100	100	98	99	0.9991	0	0.98
Quadratic	97	100	100	97.1	98.5	0.9887	0	0.97
Fine Gaussian	99	98	98	99	98.5	0.9922	0.03	0.99
Coarse Gaussian	94.1	100	100	94.3	97	0.9986	0	0.93
Nearest neighbor								
Fine KNN	97	98	98	97	97.5	0.975	0.01	0.98
Coarse KNN	74.3	100	100	79.2	87	0.9961	0	0.75
Weighted KNN	97	99	99	97	98	0.994	0	0.99
Decision tree								
Complex tree	99	98	98	99	98.5	0.9815	0.02	0.99
Medium tree	99	98	98	99	98.5	0.9815	0.02	0.99
Ensemble								
Boosted tree	100	88.9	90.2	100	94.5	0.9001	0	0.96
Bagged tree	98	99	99	98	98.5	0.9996	0.01	0.98
Subspace disc.	95	100	100	95.2	97.5	0.9986	0	0.94
Subspace KNN	99	99	99	98	98.5	0.9992	0.01	0.98

Table 2Using several machine learning classifiers, classification results were used to identify healthy set O individuals from epileptic seizure (set S, ictal intervals) patients

KS	TNR (%)	TPR (%)	NPV (%)	PPV (%)	Overall accuracy (%)	AUC
1	99	100	100	99	99.5	0.9987
2–6	98	100	100	98	99	0.9987
7–9	97	100	100	97.1	98.5	0.9989
10–11	96	100	100	96.1	98	0.9988
12	95	100	100	95.2	95.7	0.9991
13–15	93.1	100	100	93.4	96.5	0.999
16	92.1	100	100	92.5	96	0.9989
17	93.1	100	100	93.4	96.5	0.9988
18–19	92.1	100	100	92.5	96	0.9985
20	90.1	100	100	90.8	95	0.9984
25	84.2	100	100	86.1	92	0.9974
30	78.2	100	100	81.8	89	0.9977
40	64.4	100	100	73.3	82	0.9982
50	55.4	100	100	68.8	77.5	0.9981

BCL	TNR (%)	TPR (%)	NPV (%)	PPV (%)	Overall accuracy (%)	AUC
2–19	99	100	100	99	99.5	0.9986
20–23	98	100	100	98	99	0.999
> 24	97	100	100	97.1	98.5	0.9982

Table 3By utilising a support vector machine (SVM) classifiers with a linear kernel & adjusting parameters with such a fixed box constraint level (BCL) value of 1, as well as a distinct set of values for kernel scale (KS), we were able to differentiate between healthy Set O subjects and subjects who were having epileptic seizures (Set S, ictal intervals).

KS	TNR (%)	TPR (%)	NPV (%)	PPV (%)	Overall accuracy (%)	AUC
1	96	99	99	96.1	97.5	0.9982
2–3	97	100	100	97.1	98.5	0.9986
4–9	98	100	100	98	99	0.9989
10–14	97	100	100	97.1	98.5	0.9989
15–16	96	100	100	96.1	98	0.9989
17	95	100	100	95.2	97.5	0.9991
18	94.1	100	100	94.3	97	0.999
19–24	93.1	100	100	93.4	96.5	0.9991
25	92.1	100	100	92.5	96	0.9986
30	89.1	100	100	90	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	98	98	95.1	96.5	0.987

Table4 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	99	94.9	95.2	98.9	97	0.9946
2–14	98	99	99	98	98.5	0.9964
15	96	100	100	96.1	98	0.999
16	95	100	100	95.2	97.5	0.9981
17	94.1	100	100	94.3	97	0.9991

18–24	93.1	100	100	93.4	96.5	0.9991
25	92.1	100	100	92.5	96	0.9985
30	89.1	100	100	90	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	96	96.2	99	97.5	0.9943

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

Distance metric	TNR (%)	TPR (%)	NPV (%)	PPV (%)	Overall accuracy (%)	AUC
Euclidean distance	98	99	99	98	98.5	0.985
City block	99	100	100	99	99.5	0.995
Chebyshev	95	99	99	95.1	97	0.9702
Cubic	93.1	100	100	93.4	96.5	0.9653
Cosine	97	99	99	97	98	0.98
Correlation	97	99	99	97	98	0.98
Spearman	98	99	99	98	98.5	0.985
Hamming	54.5	97	94	67.6	75.5	0.7512
Jaccard	54.5	97	94	67.6	75.5	0.7512

Table 6 shows the performance evaluation results that were used to differentiate healthy set O participants from epileptic seizure (set S, ictal intervals) using a K-nearest neighbour classifier to variable distance metrics & selection criteria of number of neighbours as 1 and equal distance weight. This was done by using a K-nearest neighbour classifier with variable distance metrics.

Results

In order to determine how well the categorization worked, we used the most trustworthy machine learning classifiers that could be quickly accessed. This allowed us to examine the findings quickly and accurately. Techniques such as Support Vector Machine Kernels, Decision Trees, KNNs, & Ensemble methods were among them. The performance of the categorization was evaluated based on several different metrics, including TPR, TNR, PPV, NPV, overall accuracy, and AUC. The classifiers are tweaked with various choices and parameters to dig deeper into the performance than the traditional classification techniques utilised in the study. This allows you to check how various classifiers perform over several scales, distance measurements, learning rates, and distance weights, among other things. Performing feature extraction & selection, in conjunction with the use of a classifier, is the single most important step in conducting an efficient investigation into any matter. The characteristics needed to diagnose epileptic seizures have previously been uncovered by researchers. In this study, we made use of a number of different feature extraction methods in order to take into consideration the dynamical characteristics of extremely intricate nonlinear, spatiotemporal variations in brain signals. Humans were able to analyse the performance of its classification while using the most reliable machine learning classifiers that were easily accessible, including such “Support Vector Machine Kernels, Decision Trees, KNNs, and Ensemble methods”. This allowed us to determine the “TPR, TNR, PPV, NPV, overall accuracy, and AUC values” for the classification. This was achieved by analysing the performance of its classification in terms of TPR, TNR, PPV, NPV, and total accuracy. The genuine class is represented by the left column figures, while the anticipated class is represented by the right column figures. The healthy participants are represented by red dots, whereas the epileptic ones are represented by green dots. Furthermore, the inaccuracies are shown by the cross in both colours. The correspondence confusion matrix provides an unmistakable illustration of the performance indicators for each of these three classifiers. Following the forecast, there were three data points—numbers 2, 6, and 27—that were anticipated to be false positives. The actual class, represented by the red crosses, and the predicted class, represented by the green crosses, are both apparent with in scatter plots. The performance of the categorization for both the actual classes and the projected classes can be visually examined using scatter plots. A graphical representation of mistake that makes use of the slack variable & non-linear separation with SVM is provided in Figure 2. This illustration is being provided for your reference only.

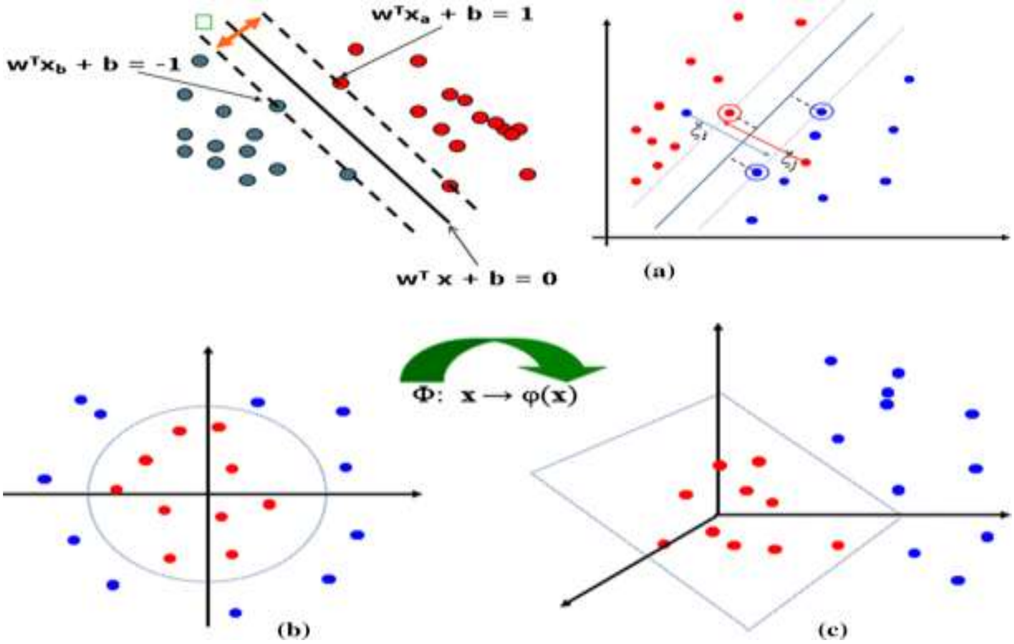

Figure 2a Error on margin using slack variable, b, c SVM non-linear separation

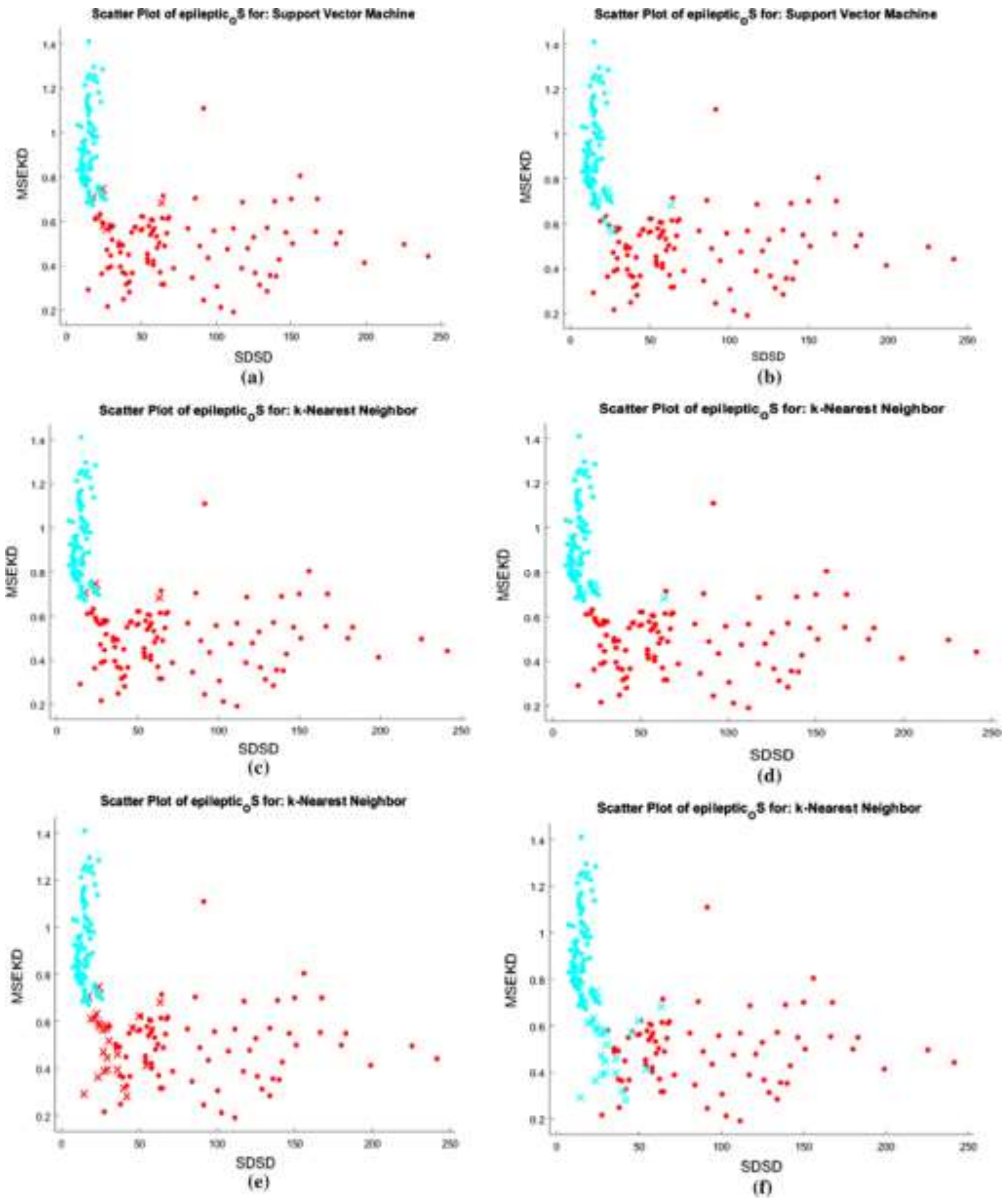

Figure 3a, b shows the SVM linear kernel (left), the true class (right), and the predicted class. Figure 3c, d shows the cosine KNN algorithm (left), the true class (right), and the anticipated class. Figure 3e, f shows the coarse KNN algorithm (left) and the true class (right)

Receiver operating curve

The measurements of the true positive rate (TPR) of healthy participants, which is also known as sensitivity, and the false positive rate (FPR) of epileptic patients, which is also called as specificity, were plotted against by the receiver operating characteristic (ROC). Those participants in the study who do not have epilepsy have a mean feature value of 1, while those participants who do have epilepsy have a mean feature value of 0. The ROC function makes use of this vector in order to visualise every sample value in relation to its specificity & sensitivity values. ROC stands for receiver operating characteristic. For the aim of categorising & demonstrating overall performance of a diagnostic system, a common practise that is followed in today's world is to use a metric that is known as the area under the receiver operating characteristic curve (ROC). On this particular plot, the TPR is displayed along the y-axis, whilst the FPR is displayed along the x-axis. This area under the curve, which reflects the square unit component, is also referred to as the AUC, which stands for area under the curve. A common acronym for "area under the curve," "AUC" stands for "area under the curve." It is possible for the number of them to be anything between 0 and 1, inclusive. The range is considered to be inclusive. When the area under the curve (AUC) is more than 0.5, there is a discernible difference between the groups. The increase in the overall quality of the diagnostic system is directly responsible for the rise in the area under the curve (AUC). To calculate the true positive rate, also referred to as the TPR, divide the total number of positive cases by the number of scenarios in which a positive outcome was anticipated. This will give you the actual positive rate. This will offer you an accurate representation of the positive rate. On the other hand, the false positive rate, which is also referred to as FPR and can be calculated by dividing the number of negative cases by the amount of instances which were anticipated to have a positive result, refers to the rate where a positive result is reported when a negative result was expected. This rate can be found by dividing the number of negative cases by amount of instances that were anticipated to have a positive result. The conclusions of the research are presented in Figure 5, and it shows that the SVM linear classifier has a separation of 0.99, while the course SVM has a separation of 0.75. Tables 1, 2, 3, 4, 5, and 6 provide tabular representations of the area under the curve (AUC) statistics for each of the several classifiers. These tables can be found below. Figure 4 is a plot of the data that illustrates a parallel coordinates plot of a data using predictions of SVM kernel function & coarse KNN classifiers vs means n SD, where n = 2, 4, 6, 8, for just a variety of extracted features. The plot shows how these classifiers perform in comparison to the means and standard deviations of the data. An parallel coordinates plot of a data is shown here as the display for this figure. This graphic is a visualisation of the data in parallel coordinates, and it is provided here.

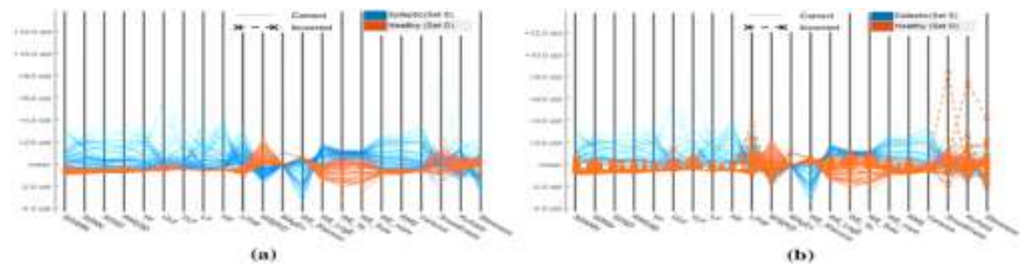

Figure 4 Shows the Detection of an Epileptic Seizure Using Several Different Feature Sets and Classifying Them Using Either an SVM Linear Kernel or a Coarse KNN

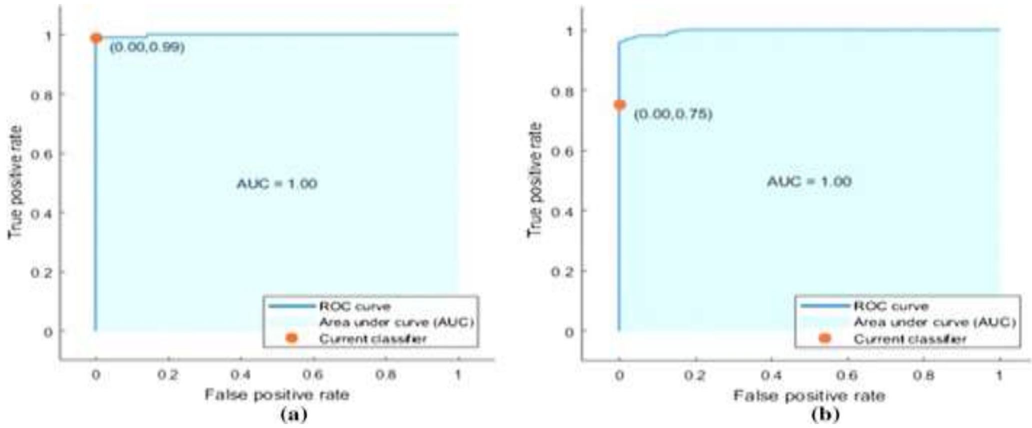

Figure 5 shows the Receiver Operating Curve (ROC) that was used to differentiate healthy participants from epileptic ones by employing an SVM linear kernel and a coarse KNN

The True Positive Rate (TPR), the True Negative Rate (TNR), the Positive Predictive Value (PPV), the Negative Predictive Value (NPV), the Overall Precision, and the Area under the Curve (AUC) have been computed to identify epileptic seizures with examples of healthier and better Set O vs Epileptic Set S subjects. These results are displayed in the ROCs, Scatter graph, and Confusion Matrix, as well as the performance assessment Tables below. The True Positive Rate The percentage of actual positive tests.

True positive rate (TPR): $TPR = \frac{TP}{TP + FN}$

True negative rate (TNR): $TNR = \frac{TN}{TN + FP}$

Positive predictive value (PPV): $PPV = \frac{TP}{TP + FP}$

Negative predictive value (NPV): $NPV = \frac{TN}{TN + FN}$

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

PPV was obtained by utilising SVM Linear kernel, that was founded on confusion matrices, confusion matrices, and scatter plots. This led to a success rate of 98 percent. When compared, PPV was obtained at a rate of 94.1 percent through the utilisation of Cosine KNN, while PPV was produced at a rate of 73.2 percent through the utilisation of Coarse KNN. The true class, which is displayed on the left, and the projected class, which is displayed on the right, are qualitatively equivalent, as evidenced by the scatterplots, which provide a graphical depiction of this factThe SVM Linear kernel have generated a total of two false positive data points, the Cosine KNN algorithm has produced a total of six erroneous positive data points, and even the Coarse KNN algorithm has produced a total of 27 false positive data points. These results are in line with the forecast. As a result of this, scatterplots can also be used to not just to determine this same connection between expected and actual classes, and also to differentiate between healthy and unhealthy circumstances based on the values of any feature. This is possible because scatterplots can be used to plot any number of features simultaneously. (Figs. 4, 5, 6).

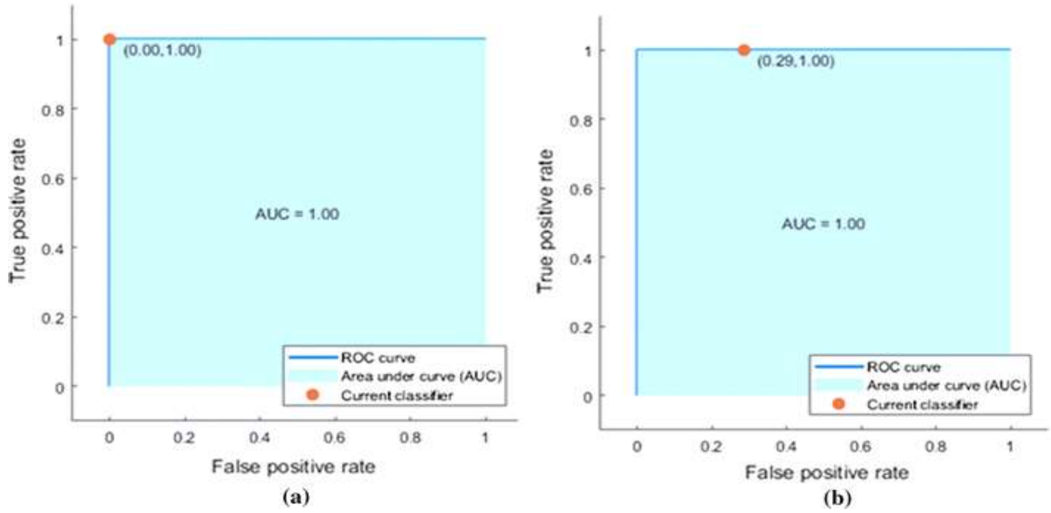

Figure 6 shows a Receiver Operating Curve (ROC) that can differentiate between epileptic subjects and postical heart rate oscillations by utilising an SVM linear kernel and a weighted KNN

The persons who are healthy and do not have epilepsy are represented by the red colour, whereas the blue colour represents the means of the epileptic patients who are in Set S. When (a) SVM linear kernel and (b) Coarse KNN are utilised, the lines represent subjects that were correctly classified, whereas the x represent data that was not properly classified. The results that were obtained and presented using Tables 1 & 2 show that the SVM linear kernel generates fewer data points that are incorrectly classified, whereas the coarse KNN generates a greater misclassification result, as seen in Figure 5b. These findings were obtained and presented in their respective tables. Additionally, the model shows that when using Coarse KNN, healthy people generate significantly more improperly classified data than epileptic participants do, whereas the SVM linear kernel generates far fewer samples that are incorrectly categorised. Figure 8a–d depicts its mean extracted features being shown against healthy & epileptic participants utilising multi-features vectors of (a) frequency domain, (b) time domain and statistics, (c) wavelet entropy, and (d) complexity based quick sample entropy features. These values were shown using the frequency domain, time domain, wavelet entropy, and complexity based quick sample entropy features. These comparisons of mean values were carried out with the help of the participants who suffered from epilepsy. According to the findings presented in Fig. 5a, b, and a–d, the mean feature values of epileptic patients are significantly higher than that of healthy individuals for every example (a–c). This is as a result of the fact that people with epilepsy are created as a result of higher levels of neurological chronic disease and higher spikes. Because the complexity of healthy individuals is greater than those of epileptic subjects, the epileptic attribute of healthy persons were shown to be significantly higher than among epileptic subjects. This outcome is consistent with findings from other research done in the past.

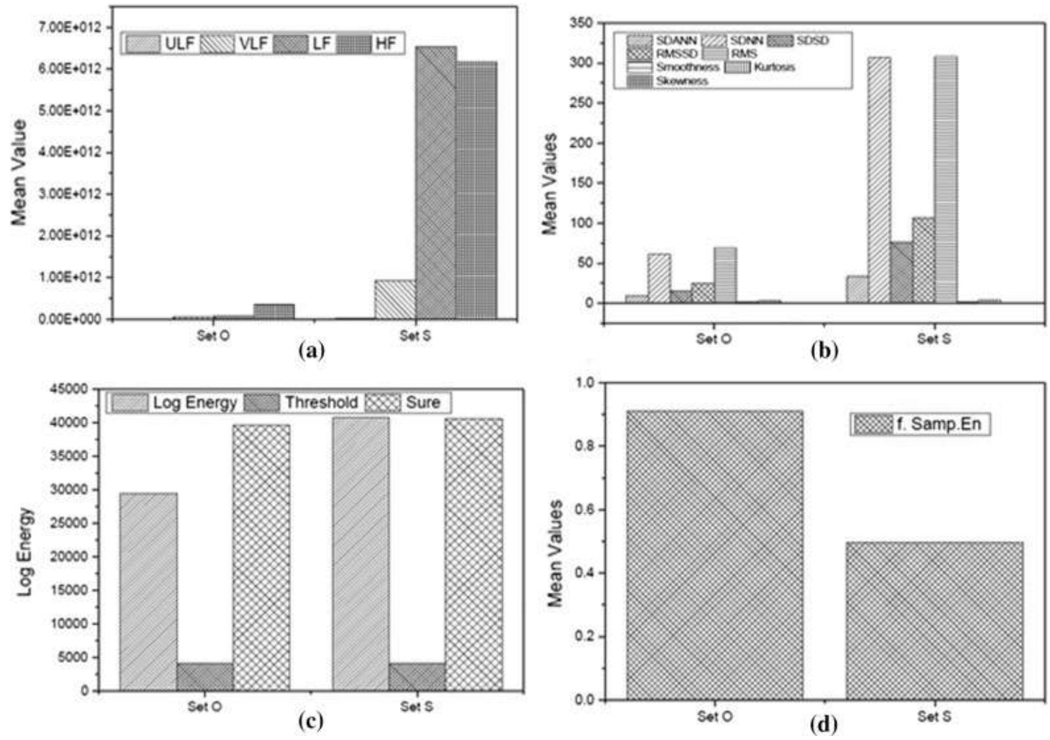

Figure 7 displays the average scores of multidomain characteristics that were retrieved from patients who were either healthy or epileptic. a characteristics of the frequency domain, b characteristics of the time domain and statistics, c characteristics of the wavelet entropy, and d characteristics of the complexity base

Table 2 displays the SVM, KNN, Decision Tree, & Ensemble classifiers along with the settings that are set for them by default. These included Support Vector Machines with parameters like linear, quadratic, fine Gaussian, and coarse Gaussian; Kernel Neural Networks with parameters like fine KNN, fine KNN, and coarse weighted KNN; & Decision Trees with parameters like complicated tree and medium tree. In order to integrate the classifier input vectors, we utilised a number of different feature extraction algorithms. When utilising SVM, the Linear, Quadratic, & Coarse Gaussian kernels achieve both 100 percent TPR and NPV, with the Fine Gaussian kernel coming in second at 98 percent. The TNR is highest for the Fine Gaussian distribution (99 percent), followed by the Linear distribution (98 percent), Quadratic distribution (97 percent), and then the Coarse Gaussian distribution (97 percent) (94.1 percent). The parameter with the highest positive predictive value (PPV) was the fine Gaussian model (99 percent), followed by the linear model (98 percent), the quadratic model (97.1 percent), and the coarse Gaussian model (94.3 percent). The overall accuracy of the SVM was 98 percent when using a linear kernel, 98 percent when using a quadratic kernel, 98 percent when using a fine Gaussian kernel, and 98 percent when using a coarse Gaussian kernel (97 percent). The linear kernel has the highest value for its area under the curve, which translates to 0.99991, followed by the fine Gaussian kernel (0.9922), the coarse Gaussian kernel (0.9987), as well as the quadratic kernel (0.9987). (0.9887). When utilising KNN in a similar fashion, Weighted KNN have been shown to have the highest accuracy (98 percent), followed by Fine KNN (97 percent) & Coarse KNN (97 percent) (87 percent). The Coarse KNN method was able to obtain the best separation (AUC = 0.9961), which was then followed by the Weighted KNN algorithm (AUC = 0.9940), and finally the Fine KNN technique (AUC = 0.9750). The table that follows offers a number of different performance ratings: Complex and medium decision trees have an accuracy of 98.5 percent, & their area under the curve is 0.9815 Complex and medium decision trees have an accuracy of 98.5 percent, & their area under the curve is 0.9815. After Bagged Tree using Subspace KNN (98.5 percent), Subspace Discriminant (97.5 percent), and Boosted Tree (97.5 percent), Bagged Tree using Subspace KNN (98.5 percent) scored the greatest Ensemble accuracy (94.5 percent). The AUC for Boosted Tree is 0.9001, but the AUC for Bagged Tree is 0.9996, and its Subspace KNN and Subspace Discriminant are both 0.9992 and 0.9986 respectively. Tables 2 and 7 present the confidence intervals for such area under the curve, which can be found in the previous table.

Classifier	**TNR (%)**	**TPR (%)**	**NPV (%)**	**PPV (%)**	**Overall accuracy (%)**	**AUC**	**95% CI**	
							LB	**UB**
Support vector machine (SVM)								
Linear	100	100	100	100	100	1	0	1
Quadratic	100	100	100	100	100	1	0	1
Fine Gaussian	100	97.09	54.14	100	97.2	1	0.43	1
Coarse Gaussian	100	99.01	85.71	100	99.1	1	0.14	1
Nearest neighbor								
Fine KNN	100	100	100	100	100	1	0	1
Coarse KNN	100	99.01	85.71	100	99.1	1	0.16	1
Weighted KNN	100	98.04	71.43	100	98.1	1	0.29	1
Decision tree								
Complex tree	100	100	100	100	100	1	0	1
Medium tree	100	100	100	100	100	1	0	1
Ensemble								
Bagged tree	100	100	100	100	100	1	0	1
Subspace disc.	100	100	100	100	100	1	0	1
Subspace KNN	100	100	100	100	100	1	0	1

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

Table 3 contains a wide variety of kernel scales & box limitations, as well as the findings of such an evaluation of SVM's performance using a linear kernel. To begin, we utilised a number of distinct kernel scales, with the restriction that we could only use one box. Table 3 demonstrates that there is a significant gap between different kernel scales in terms of the performance measures like TNR, PPV, overall accuracy, and AUC. These values can be found below. This Kernel's accuracy went from 99 to 99.5 percent with KS = 1 and BCL = 1 after certain tweaks were made to the Linear Support Vector Machine's (SVM) default kernel size and box constraint levels. After other kernel scales (KS 2–6), which have an accuracy of 99 percent, comes the accuracy of 98.5 percent (KS 7–9), then by the accuracy of 98 percent (KS 10–11), and so on. At KS 55, we achieved a score of 77.5 percent. Both TPR & NPV remained unchanged at 100%. Both TNR and PPV are equally accurate in terms of the big picture. It was revealed that KS 12 had the greatest level of separation (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), and KS 18–19 (AUC = 0.9985). The AUC can be used to discriminate between healthy individuals and those that are experiencing various illnesses. There are a multitude of different scales that can help differentiate between participants, as seen by the AUC values at the various KS. These deeper, more meaningful characteristics provide insight on the significantly enhanced performance. On our second attempt, we corrected the KS and tested a few other BCL. In general, the accuracy of the BCL 2–19 has been the highest (99.5 percent), followed by the accuracy of the BCL 20–23 (99 percent), and then by the accuracy of the BCL > 24 (98 percent) (98.5 percent). The area under the curve for BCL 20–23 had the greatest value (0.9990), followed by the area under the curve for BCL 2–19 (0.9986), and so on.

In order to identify between people who were healthy and those who had epilepsy, the SVM Quadratic Kernel at Multiple Kernel Scales and the BCL were utilised. Again, the KS 4–9 range had the maximum accuracy, coming in at 99 percent, in contrast to the 88.5 percent achieved by employing fixed scales & BCL. After this, the Key Stages 2–3, 10–16, and Key Stages 1–17 came into play, respectively. The greatest amount of differentiation was observed between KS 17, 19–24 (AUC = 0.9991), which was then followed by KS 18 (AUC = 0.9990), KS 4–16 (AUC = 0.9989), KS 2–3, 25 (AUC = 0.9986), etc. Our findings indicated that the accuracy was 96.5 percent if KS is set to 1 and BCL has been more than 2. Additionally, the PPV was 95.1 percent, the TNR was 95 percent, the TPR was 95 percent, and the NPV was 98 percent. AUC 0.9870

The SVM with a Gaussian kernel is illustrated in Table 5 with the KS & BCL values. In this situation, the accuracy for KS 2–14 was 98.5 percent, the accuracy for KS 15 is 98.5 percent, the accuracy for KS 16 is 97.5 percent, and so on. In a manner analogous to that of KS 17–24, KS 15 have had highest AUC, which was 0.9990, followed by KS 25 with 0.9985 and KS 30 with 0.9982. As KS increased, TPR and PPV both decreased, although TNR and NPV both climbed. By adjusting KS to 1 & increasing BCL to be more than 2, we were able to get an accuracy of 97.5 percent, a TNR of 99 percent, a TPR of 96 percent, a PPV of 99 percent, an NPV of 96.2 percent, and an AUC of 0.9943.

Table 6 contains the findings that can be obtained from applying KNN's performance analysis to a number of different distance measures. The KNN algorithm combined with the City Block distance metric seemed to have the highest accuracy at 99.5 percent, followed by 98.5 percent when combined with the Euclidean Distance, Spearman; 98 percent when combined with Cosine & Correlation; 97 percent when combined with Chebyshev; 96.5 percent when combined with Cubic; and 75.5 percent when combined with the Hamming and Jaccard distance metric. It was discovered that the area under the curve (AUC) for the City Block distance measure was 0.9950, that the Euclidean Distance & Spearman distance measure was 0.9850, that the Cosine & Correlation distance measure was 0.9800, that the Chebyshev distance measure had an AUC of 0.9702, that the Cubic distance measure had an AUC of 0.9653, and that the Hamming & Jaccard distance measure had an AUC of 0.7 Table 6 displays the TNR, TPR, NPV, and PPV for your reference.

The support vector machine with linear and quadratic kernels, fine KNN, or ensemble techniques like bagged tree, subspace discriminate, & subspace KNN offer the highest possible performance, with a TPR, TNR, PPV, NPV, accuracy, and AUC of 1.00 and 100 percent respectively. TNR (100 percent) TPR (99.01 percent) NPV (85.71 percent) PPV (100 percent) AUC (1.00). The accuracy of the SVM fine Gaussian is 97.2 percent, its TNR is 100 percent, its TPR is 97.9 percent, its NPV is 54.1 percent, and its PPV and AUC are both 100% percent respectively (1.00). However, coarse KNN gave a TPR of 99.01 percent, NPV of 85.71 percent, and accuracy of 99.1 percent, whereas fine, coarse, & weighted KNN all gave 100 percent for both TNR and PPV. A TPR of 98.04 percent, NPV of 71.43 percent, & accuracy of 99.1 percent were obtained by the use of weighted KNN. (98.1 percent). Table 7 presents the confidence intervals for each classifier based on a percentage of 95.

Figure 8 depicts the accuracy of the KNN algorithm with different Neighbors & distance weights. Greatest accuracy is achieved with all distance weights (equal, inverse, & squared inverse) with lower numbers of Neighbors (1 and 4), but this accuracy decreases even as number of Neighbors increases. The maximum accuracy is achieved with all distance weights (98.5 percent) (Fig. 9). The accuracy was as follows when using the same distance weight: N8 (97.5 percent), N10 (97.5 percent), N20 (96 percent), N30 (96 percent), N50 (93.5 percent), and N100 (100 percent) (86.5 percent). The accuracy was as follows when using the inverse distance weight: N8 (98 percent), N10 (97.5 percent), N20 (96.5 percent), N30 (96 percent), N50 (95 percent), and N100 (100 percent) (90 percent). In a similar manner, by utilising the squared inverse distance weight, i were able to arrive at N8 (98.5 percent), N10 (98.5 percent), N20 (96.5 percent), N30 (96.5 percent), N50 (96 percent), and N100 (100 percent) (98.5 percent). (96 percent).

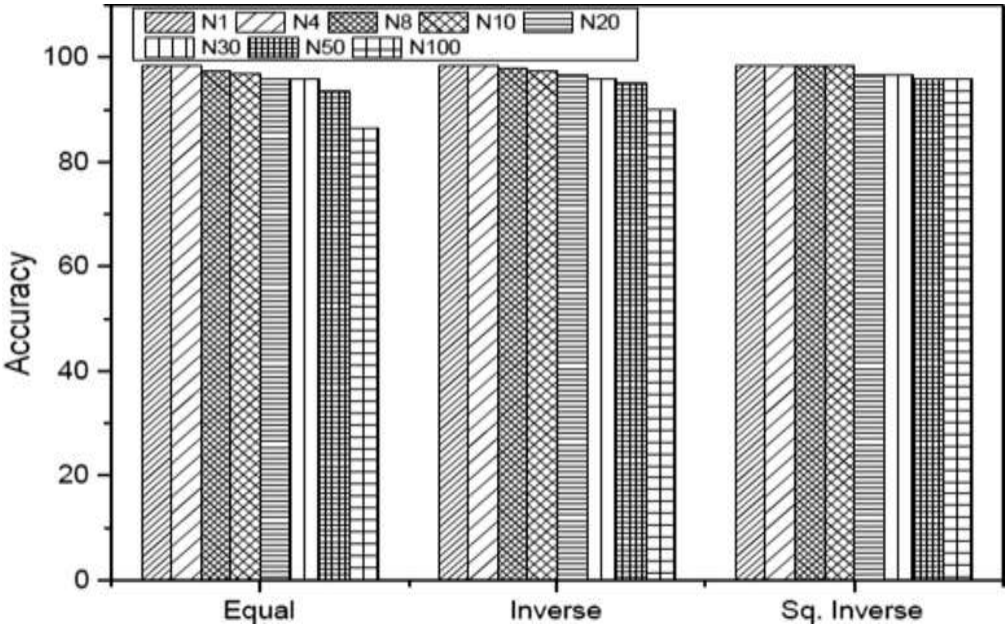

Table 8: Overall Accuracy Obtained to Distinguish Healthy Set O Subjects with Epileptic Seizure (Set S, Ictal Intervals) Using K-Nearest Neighbor Classifier Utilizing Euclidean Distance With Varying Numbers of Neighbors (1, 4, 8, 10, 20, 30, 50, and 100) And Distance Weight

Conclusion

Epilepsy is responsible for the deaths of millions each and every year. For diagnostic and prognostic analysis, such complex and nonstationary datasets call for multidomain feature extraction strategies, in addition to robust machine learning classifiers. We provided multidomain features based on the characteristics of these signals. These features included time-domain, frequency-domain, statistical, & complexity measures, and also wavelet-based entropy assessments. They make use of Support Vector Machine Kernels, Decision Trees, K-nearest Neighbors, & Ensemble Classifiers with increased parametrization in order to recognise epileptic convulsions. This automated method has the potential to support the efforts of medical personnel with saving millions of lives every year. The processes of categorization and detection both made use of this hybrid feature set in their respective approaches. SVM classifiers have a propensity to produce superior results over a wide range of kernel sizes, parameterizations, & learning rates; however, KNN classifiers have a propensity to perform better when dealing with a variety of distance metrics and Neighbors. When compared to the approaches that came before them, the procedures that possess multidomain features yield results that are of a higher quality. When attempting to discern among postictal heart beat oscillations & epileptic seizures with ictal intervals, the use of robust machine learning classifiers that make use of multimodal feature extraction has proven to yield the best possible results.

References

Adeli, H., Zhou, Z., &Dadmehr, N. (2003). Analysis of EEG records in an epileptic patient using wavelet transform. *Journal of neuroscience methods*, *123*(1), 69-87.

Andrzejak, R. G., Lehnertz, K., Mormann, F., Rieke, C., David, P., &Elger, C. E. (2001). PART 1 REGULAR ARTICLES-Biological physics-Indications of nonlinear deterministic and finite-dimensional structures in time series of Brain electrical activity: Dependence on recording region and. *Physical Review-Section E-Statistical Physics Plasma Fluids Related Interdiscpl Topics*, *64*(6).

Ayoubian, L., Lacoma, H., &Gotman, J. (2013). Automatic seizure detection in SEEG using high frequency activities in wavelet domain. *Medical engineering & physics*, *35*(3), 319-328.

Bashar, S. K., Hassan, A. R., & Bhuiyan, M. I. H. (2015, August). Identification of motor imagery movements from EEG signals using dual tree complex wavelet transform. In *2015 international conference on advances in computing, communications and informatics (ICACCI)* (pp. 290-296). IEEE.

Bashar, S. K., Hassan, A. R., & Bhuiyan, M. I. H. (2015, December). Motor imagery movements classification using multivariate EMD and short time Fourier transform. In *2015 Annual IEEE India Conference (INDICON)* (pp. 1-6). IEEE.

Fu, K., Qu, J., Chai, Y., & Zou, T. (2015). Hilbert marginal spectrum analysis for automatic seizure detection in EEG signals. *Biomedical Signal Processing and Control*, *18*, 179-185.

Ghosh-Dastidar, S., Adeli, H., &Dadmehr, N. (2007). Mixed-band wavelet-chaos-neural network methodology for epilepsy and epileptic seizure detection. *IEEE transactions on biomedical engineering*, *54*(9), 1545-1551.

Gotman, J. (1982). Automatic recognition of epileptic seizures in the EEG. *Electroencephalography and clinical Neurophysiology*, *54*(5), 530-540.

Gotman, J., &Gloor, P. (1976). Automatic recognition and quantification of interictal epileptic activity in the human scalp EEG. *Electroencephalography and clinical neurophysiology*, *41*(5), 513-529.

Guo, L., Rivero, D., &Pazos, A. (2010). Epileptic seizure detection using multiwavelet transform based approximate entropy and artificial neural networks. *Journal of neuroscience methods*, *193*(1), 156-163.

Guo, L., Rivero, D., Dorado, J., Munteanu, C. R., &Pazos, A. (2011). Automatic feature extraction using genetic programming: An application to epileptic EEG classification. *Expert Systems with Applications*, *38*(8), 10425-10436.

Hassan, A. R., & Bhuiyan, M. I. H. (2015, December). Dual tree complex wavelet transform for sleep state identification from single channel electroencephalogram. In *2015 IEEE International Conference on Telecommunications and Photonics (ICTP)* (pp. 1-5). IEEE.

Hassan, A. R., & Bhuiyan, M. I. H. (2017). An automated method for sleep staging from EEG signals using normal inverse Gaussian parameters and adaptive boosting. *Neurocomputing*, *219*, 76-87.

Hassan, A. R., &Haque, M. A. (2016). Computer-aided obstructive sleep apnea screening from single-lead electrocardiogram using statistical and spectral features and bootstrap aggregating. *Biocybernetics and Biomedical Engineering*, *36*(1), 256-266.

Hassan, A. R., &Haque, M. A. (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, A. R., Siuly, S., & Zhang, Y. (2016). Epileptic seizure detection in EEG signals using tunable-Q factor wavelet transform and bootstrap aggregating. *Computer methods and programs in biomedicine*, *137*, 247-259.

Iscan, Z., Dokur, Z., &Demiralp, T. (2011). Classification of electroencephalogram signals with combined time and frequency features. *Expert Systems with Applications*, *38*(8), 10499-10505.

Kang, J. H., Chung, Y. G., & Kim, S. P. (2015). An efficient detection of epileptic seizure by differentiation and spectral analysis of electroencephalograms. *Computers in biology and medicine*, *66*, 352-356.

Kannathal, N., Choo, M. L., Acharya, U. R., &Sadasivan, P. K. (2005). Entropies for detection of epilepsy in EEG. *Computer methods and programs in biomedicine*, *80*(3), 187-194.

Mormann, F., Andrzejak, R. G., Elger, C. E., &Lehnertz, K. (2007). Seizure prediction: the long and winding road. *Brain*, *130*(2), 314-333.

Nigam, V. P., &Graupe, D. (2004). A neural-network-based detection of epilepsy. *Neurological research*, *26*(1), 55-60.

Ocak, H. (2009). Automatic detection of epileptic seizures in EEG using discrete wavelet transform and approximate entropy. *Expert Systems with Applications*, *36*(2), 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 Systems with Applications*, *38*(10), 13475-13481.

Schwenker, F., &Trentin, E. (2014). Pattern classification and clustering: A review of partially supervised learning approaches. *Pattern Recognition Letters*, *37*, 4-14.

Stochholm, A., Mikkelsen, K., &Kidmose, P. (2016, August). Automatic sleep stage classification using ear-EEG. In *2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)* (pp. 4751-4754). IEEE.

Subasi, A. (2007). EEG signal classification using wavelet feature extraction and a mixture of expert model. *Expert Systems with Applications*, *32*(4), 1084-1093.

Tzallas, A. T., Tsipouras, M. G., & Fotiadis, D. I. (2007). Automatic seizure detection based on time-frequency analysis and artificial neural networks. *Computational intelligence and neuroscience*, *2007*.

Wang, R., Kwong, S., Wang, X. Z., & Jiang, Q. (2014). Segment based decision tree induction with continuous valued attributes. *IEEE transactions on cybernetics*, *45*(7), 1262-1275.

Zhang, P., Gao, B. J., Zhu, X., & Guo, L. (2011, December). Enabling fast lazy learning for data streams. In *2011 IEEE 11th International Conference on Data Mining* (pp. 932-941). IEEE.