Weighted Majority Voting Based Ensemble of Classifiers Using Different Machine Learning Techniques for Classification of EEG Signal to Detect Epileptic Seizure

Sandeep Kumar Satapathy
Department of Computer Science and Engineering
Siksha 'O' Anusandhan University, Khandagiri, Bhubaneswar-751030, Odisha, India
E-mail: sandeepkumar04@gmail.com

Alok Kumar Jagadev
School of Computer Engineering
KIIT University, Bhubaneswar-751024, Odisha, India
E-mail: alok.jagadev@gmail.com

Satchidananda Dehuri
Department of Information and Communication Technology
Fakir Mohan University, Vyasa Vihar-756019, Balasore, Odisha, India
E-mail: satchi.lapa@gmail.com

Keywords: EEG signal, epilepsy, classification, machine learning

Received: January 6, 2017

Electroencephalogram (EEG) signal is a miniature amount of electrical flow in a human brain that holds and controls the entire body. It is very difficult to understand these non-linear and non-stationary electrical flows through naked eye in the time domain. In specific, epilepsy seizures occur irregularly and un-predictively while recording EEG signal. Therefore, it demands a semi-automatic tool in the framework of machine learning to understand these signals in general and to predict epilepsy seizure in specific. With this motivation, for wide and in-depth understanding of the EEG signal to detect epileptic seizure, this paper focus on the study of EEG signal through machine learning approaches. Neural networks and support vector machines (SVM) basically two fundamental components of machine learning techniques are the primary focus of this paper for classification of EEG signals to label epilepsy patients. The neural networks like multi-layer perceptron, probabilistic neural network, radial basis function neural networks, and recurrent neural networks are taken into consideration for empirical analysis on EEG signal to detect epilepsy seizure. Furthermore, for multi-layer neural networks different propagation training algorithms have been studied such as back-propagation, resilient-propagation, and quick-propagation. For SVM, several kernel methods were studied such as linear, polynomial, and RBF during empirical analysis. Finally, the study confirms with the present setting that, in all cases recurrent neural network performs poorly for the prepared epilepsy data. However, SVM and probabilistic neural networks are quite effective and competitive. Hence to strengthen the poorly performing classifier, this work makes an extension over individual learners by ensembling classifier models based on weighted majority voting.

Povzetek: Sistem s pomočjo strojnega učenja iz EEG signalov zazna epileptični napad.

1 Introduction

Epilepsy is a persistent disorder of mental ability that has an abnormal EEG signal flow [1], which manifests in the disoriented human behaviour. In this world, around 40 to 50 million people are mostly affected by this disease [2]. Many people also call it as *fits* that causes loss of memory and interruption in consciousness, strange sensations, and significant alteration in emotions and behaviour. Research related to the epilepsy disease is basically used for differentiating between Ictal (seizure period) and Interictal (period between seizures) EEG signals. Hence, the transition from preictal to ictal state for an epileptic seizure contains a gradual change from a chaotic to ordered wave forms. Moreover, the amplitude of the spikes does not necessarily signify the harshness of seizures [2].

The difference between the seizure and the common artifact is quite easy to recognize where generally the seizures within EEG measurement [3] have a prominent spiky, repetitive, transient, or noise-like pattern. Hence, unlike other general signals, for an untrained observer, the EEG signal is quite difficult for understanding and analysis. The recording of these signals is mostly done by the help of a set of electrodes placed on the scalp using 10 to 20 electrode placement systems. The system incorporates the electrodes, which are placed with a specific name based on specific parts of the brain, e.g., Frontal Lobe (F), Temporal Lobe (T), etc. These naming and placement schemes have been discussed in more details in [3].

For the facilitation and effective diagnosis of the epilepsy, several neuro-imaging techniques such as functional magnetic resonance imaging (fMRI) and position emission tomography (PET) are used. An epileptic seizure can be characterized by paroxysmal occurrence of synchronous oscillation. This impersonation can be separated into two categories depending on the extent of involvement in different brain regions such as focal or partial and generalized seizures [4]. Focal seizure also known as *epileptic foci* are generated at specific sphere in the brain. In contrast to this, generalized seizures occur in most parts of the brain.

A careful analysis and diagnosis of EEG signals for detecting epileptic seizure in the human brain usually contribute to a substantial insight and support to the medical science. Thus, EEG is quite a beneficial as well as a cost effective way for the study of epilepsy disease. For generalized seizure, the duration of seizure can be easily detected by naked eyes whereas it is very difficult to recognize intervals during focal epilepsy.

Classification is the most useful and functional technique [5] for properly detecting the epileptic seizures in EEG signals. Classification being a data mining technique is generally used for pattern recognition [6]. Other than that, it is used to predict a group membership for unknown data instances. Hence, by designing a classifier model using different machine learning approaches, we can identify epileptic seizures in EEG brain signal. Pre-processing is considered as one of the necessary task to get it into a proper feature set format, even before considering the classification of raw EEG signal. Generally, the data sample of EEG is not linearly separable. Thus, to obtain non-linear discriminating function for classification, we are using machine learning techniques. Moreover, the case of limiting our focus on machine learning approaches is because of their capability and efficiency for smooth approximation and pattern recognition. However, there are learners who are performing very poorly; hence this work makes an extension over individual learners by ensembling classifier models based on weighted majority voting.

In this analytical study, a publicly available EEG dataset have been considered that is related to epilepsy for all experimental evaluations. Based on this there are mainly two phases of the epileptic seizure detection process that are carried out. The first phase is to analyse the EEG signal and convert it into a set of samples with set of features. The second phase is to classify the already processed data into different classes such as epilepsy or normal.

The rest of the subdivisions of this paper are organized as follows. Section 2 describes the recording and pre-processing of EEG signals through discrete wavelet transform. Some classification methods based on machine learning techniques are described in Section 3. Section 4 discusses the ensemble of classifiers. In Section 5, the detail of empirical work and analysis of results obtained by different machine learning models and ensemble of classifiers. Section 6 draws the conclusions and suggests possibilities for future work.

2 Methods for dataset preparation

In the present work we have collected data from [7] which is a publicly available database [8] related to diagnosis of epilepsy. This resource provides five sets of EEG signals. Each set contains reading of 100 single channel EEG segments of 23.6 seconds duration each. These five sets are described as follows. Datasets A and B are considered from five healthy subjects using a standardized electrode placement system. Set A contains signals from subjects in a slowed down state with eyes open. Set B also contains signal same as A but ones with the eyes closed. The data sets C, D and E are recorded from epileptic subjects through intracranial electrodes for interictal and ictal epileptic activities. Set D contains segments recorded from within the epileptogenic zone during seizure free interval. Set C also contains segments recorded during a seizure free interval from the hippocampal formation of the opposite hemisphere of the brain. The set E only contains segments that are recorded during seizure activity. All signals are recorded through the 128 channel amplifier system. Each set contains 100 single channel EEG data. In all there are 500 different single channel EEG data. In Subsection 2.1, we illustrate how to crack these signals using discrete wavelet transform [9] and prepare several statistical features to form a proper sample feature dataset.

2.1 Wavelet transform

This is a modern signal analysis technique which overcomes the limitations of other transformation techniques. Other transformation methods may include Fast Fourier Transform (FFT), Short Time Fourier Transform (STFT), etc. The major restrictions of these techniques are the analysis limits to stationary signals. These are not effective for analysis of transient signals such as EEG signal. Transient in the sense the frequency is changing rapidly with respect to time. Then, with the help of wavelet coefficients [10] we can analyse transient signals easily and also efficiently. Wavelet transform can be of two types: Continuous Wavelet Transform (CWT) [11] and Discrete Wavelet Transform (DWT) [12, 13].

2.1.1 Continuous wavelet transform

It is defined as:

$$CWT(a,b) = \int_{-\infty}^{\infty} x(t).\varphi_{a,b}^{*}(t)dt, \quad (1)$$

where, $x(t)$ represents the original signal, a and b represents the scaling factor and translation along the time axis, respectively. The * symbol denotes the complex conjugation and $\varphi^*_{a,b}$ is computed by scaling the wavelet at time band scale a.

$$\varphi^*_{a,b}(t) = \frac{1}{\sqrt{|a|}}\varphi\left(\frac{t-b}{a}\right), \quad (2)$$

where, $\varphi^*_{a,b}(t)$ stands for the mother wavelet. In CWT it is presumed that the scaling and translation parameters a and b changes continuously. But the main disadvantage of CWT is the calculation of wavelet coefficients for every possible scale can result in a large amount of data. It can surmount with the help of DWT.

2.1.2 Discrete wavelet transform

It is almost same as CWT except that the value of a and b does not change continuously. It can be defined as:

$$DWT = \frac{1}{\sqrt{|2^p|}}\int_{-\infty}^{\infty} x(t)\varphi\left(\frac{t-2^p q}{2^p}\right)dt, \quad (3)$$

where a and b of CWT are replaced in DWT by 2^p&2^q respectively.

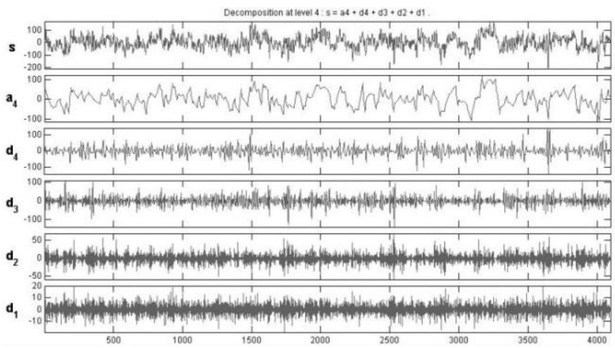

Figure 1: Single channel EEG signal decomposition of set A using db-2 up to level 4.

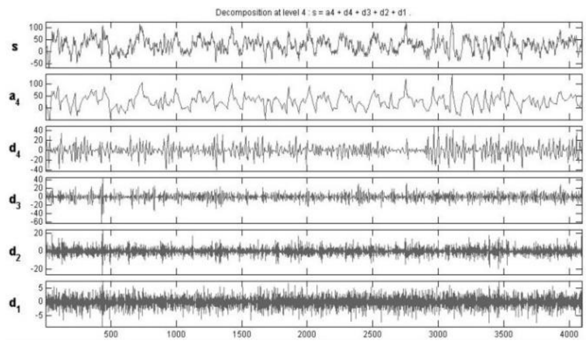

Figure 2: Single channel EEG signal decomposition of set D using db-2 up to level 4.

It is a transformation technique that provides a new data representation which can spread to multiple scales. Therefore, the analysis of transforming signal can be performed at a multiple resolution scale. DWT is performed by successively passing the signal through a series of high pass and low pass filters producing a set of detail and approximation coefficients. This generates a decomposing tree known as Mallat's decomposition tree. In this analytical work, the raw EEG signals that have been picked up from web resources is decomposed using DWT [13] available as a toolbox in MATLAB. This signal is decomposed using the Daubechis Wavelet function of order 2 up to 4 levels [12]. Thus, it produces a series of wavelet coefficient like four detailed coefficients (D1, D2, D3, and D4) and an approximation signal (A4). Figures 1, 2, and 3 provides a snapshot of this decomposition of a single channel EEG recording from set A, D, and E respectively.

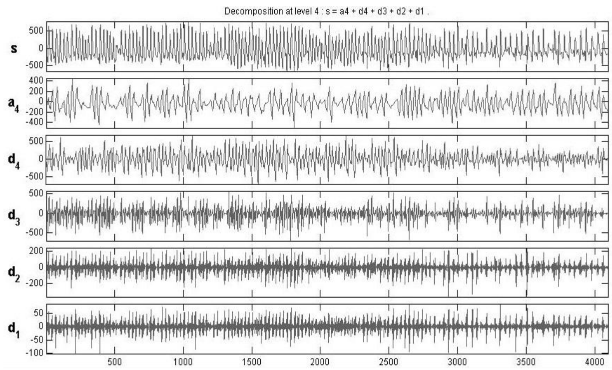

Figure 3: Single channel EEG signal decomposition of set E using db-2 up to level 4.

Later, on this decomposition some of the statistical features of many have been extracted from the signals such as Minimum (MIN), Maximum (MAX), MEAN, and Standard Deviation (SD). Figure 4 is a sample output of the MATLAB toolbox showing different features of a single channel EEG recording from set A and set E. The same procedure can be followed for all other EEG recordings to make a perfect set. So, after this level, we are ready with a sample feature dataset of order 500 by 20 matrixes as shown in Table 1. Then it can be further used for classification tasks.

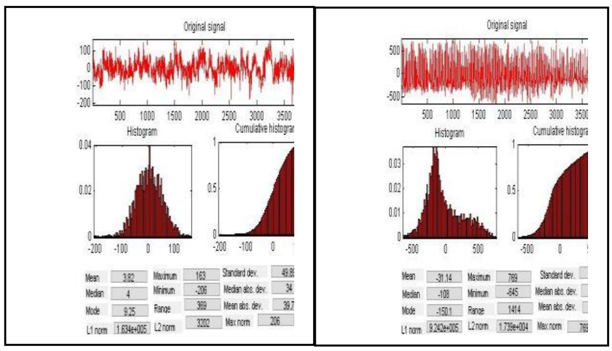

Figure 4: Statistical features extraction from signals after decomposition.

In addition to DWT, there are other feature extraction techniques [5] that can also be used successfully to extract features from the raw EEG signal. These techniques may include Wavelet Packet Decomposition (WPD), Principal Component Analysis (PCA), Lyapunov Exponent, ANNOVA test, etc.

Table 1: Structure and dimension of dataset for EEG signal classification.

Seizure Detection Sets	Size of Sample	Class 0	Class 1
Set1- (A & E)	200x20	100x20	100x20
Set 2- (D & E)	200x20	100x20	100x20
Set 3- (A+D & E)	300x20	200x20	100x20

3 Machine learning classifiers

Machine learning (ML) is a set of computerized techniques, which focus to automatically learn to recognize complex patterns and make intelligent decisions based on data. ML has proven its ability to uncover hidden information present in large complex datasets. Using ML, it is possible to cluster similar data, classify, or to find association among various features [14, 15]. In the context of EEG signal analysis, ML is the application of algorithms for extracting patterns from EEG signals [16]. However, there are other steps also carried out e.g., data cleaning &pre-processing, data reduction & projection, incorporation of prior knowledge, proper validation and interpretation of results while analysing EEG signals. EEG analysis has number of challenges which make it suitable for machine learning techniques [16].

- EEG comes in large databases.
- EEG recordings are very noisy.
- EEG signals have large temporal variance.

Some popular machine learning approaches are neural networks, evolutionary algorithms, fuzzy theory, and probabilistic learning. In this analytical work, our focus is restricted with neural networks, its variants and support vector machines for classification EEG signals.

3.1 Multilayer perceptron neural network (MLPNN)

Artificial neural network simulates the operation of a neural network of the human brain and solves a problem. Generally, single layer Perceptron neural networks are sufficient for solving linear problems, but nowadays the most commonly employed technique for solving nonlinear problems is Multilayer Perceptron Neural Network (MLPNN) [17]. It can hold various layers such as one input and one output layer along with at least one hidden layer. There are connections between different layers for data transmission. The connections are generally weighted edges to add some extra information's to the data and it can be propagated through different activation functions.

The heart of designing an MLPNN is the training of network for learning the behaviour of input-output patterns. In this work, we have designed an MLPNN with the help of a Java Encog framework. This network is trained with the help of three popular training algorithms such as Back- propagation (BP) [18], Resilient Propagation (RPROP) [19], and Manhattan Update Rule (MUR).

Back-propagation training algorithm [5, 19, and 20] is different from other algorithms in terms of the weight updating strategies. In back propagation [21, 22, 23], generally weight is updated by the equation (4) [24, 25, 26].

$$w_{ij}(k+1) = w_{ij}(k) + \Delta w_{ij}(k), \tag{4}$$

where in regular gradient decent

$$\Delta w_{ij}(k) = -\eta \frac{\partial E}{\partial w_{ij}}(k) \tag{5}$$

with a momentum term

$$\Delta w_{ij}(k) = -\eta \frac{\partial E}{\partial w_{ij}}(k) + \mu \Delta w_{ij}(k-1) \tag{6}$$

Resilient propagation [19] is a supervised training algorithm for feed forward neural network. Instead of magnitude, it takes into account only the sign of the partial derivative, or gradient decent and acts independently on each weight. The advantage of RPROP algorithm is that it needs no setting of parameters before applying it. The weight updating is done according to the equation (7). Equation 4 is same for the RPROP for weight update.

$$\Delta w_{ij} = \begin{cases} +\Delta_{ij}, if \frac{\partial E}{\partial w_{ij}}(k) > 0, \\ +\Delta_{ij}, if \frac{\partial E}{\partial w_{ij}}(k) < 0, \\ 0, Otherwise \end{cases} \tag{7}$$

$$\Delta_{ij} = \begin{cases} \eta_+ * \Delta_{ij}(k-1), S_{ij} > 0, \\ \eta_- * \Delta_{ij}(k-1), S_{ij} < 0, \\ \Delta_{ij}(k-1), Otherwise. \end{cases} \tag{8}$$

where, $S_{ij} = \frac{\partial E}{\partial w_{ij}}(k-1) * \frac{\partial E}{\partial w_{ij}}(k)$ and $\eta_+ = 1.2$ and $\eta_- = 0.5$.

Manhattan update rule also works similar to RPROP and only uses the sign of the gradient and magnitude is discarded. If the magnitude is zero, then no change is made to the weight or threshold value. If the sign is positive, then the weight or threshold value is increased by a specific amount defined by a constant. If the sign is negative, then the weight or the threshold value decreases by a specific amount defined by a constant. This constant must be provided to the training algorithm as a parameter.

3.2 Variants of neural network

In addition to MLPNN, many different types of neural networks have been developed over the year for solving problems with varying complexities of pattern classification. Some of these includes: Recurrent Neural Network (RNN) [41], Probabilistic Neural Network (PNN) [42], and Radial Basis Function Neural Network (RBFNN) [43].

3.2.1 Recurrent neural network

RNN [44] is a special type of artificial neural network having a fundamental feature is that the network contains at least one feedback connection [45], so that activation can flow round in a loop. This feature enables the network to do temporal processing and learn the patterns. The most important common features shared by all types of RNN [46, 47] are, they incorporate some form of Multilayer Perceptron as sub-system. They implement the non-linear capability of MLPNN [48, 49] with some form of memory. In this research work the ANN architecture, we have implemented for modelling and classifying is the Elman Recurrent Neural Network (ERNN). It was originally developed by Jeffrey Elman in 1990. The Back-Propagation through time (BPTT) learning algorithm is used for training [50, 51], which is an extension of Back-propagation that performs gradient decent on a complete unfolded network.

If a network training sequence starts on time t_0 and ends at time t_1, the total cost function can be calculated as:

$$E_{total}(t_0, t_1) = \sum_{t=t_0}^{t_1} E_{\frac{sse}{ce}}(t), \tag{9}$$

and the gradient decent weight update can be calculated as:

$$\Delta w_{ij} = -\eta \sum_{t=t_0}^{t_1} \frac{\partial E_{\frac{sse}{ce}}(t)}{\partial w_{ij}}. \tag{10}$$

3.2.2 Probabilistic neural network

PNN was first proposed by Specht in 1990. It is a classifier that maps input patterns in a number of class levels. It can be forced into a more general function approximator. This network is organized into a multilayer feed forward network with input layer, pattern layer, summation layer, and the output layer. PNN [52] is an implementation of a statistical algorithm called kernel discriminant analysis. The advantages of PNN are like; it has a faster training process as compared to Back-Propagation. Also, there are no local minima issues. It has a guaranteed coverage to an optimal classifier as the size of the training set increases. But it has few disadvantages like slow execution of the network because of several layers and heavy memory requirements, etc.

In PNN [52] a Probability Distribution Function (PDF) is computed for each population. An unknown sample *s* belongs to a class *p* if,

$$PDF_p(s) > PDF_q(s) \forall p \neq q, \tag{11}$$

where, $\mathrm{PDF}_k(s)$ is the PDF for class *k*.

Other parameters used are Prior Probability - *h*, Misclassification Cost – *c*, so the classification decision becomes,

$$h_p c_p PDF_p(s) > h_q c_q PDF_q(s) \forall p \neq q \tag{12}$$

PDF for a single sample can be calculated by using the formula,

$$PDF_k(s) = \frac{1}{\sigma} W\left(\frac{s - s_k}{\sigma}\right), \tag{13}$$

where s – Input (unknown), s_k - k^{th} sample, W- weighting function, σ - smoothing parameter. PDF for a single population can be calculated by taking the average of PDF of *n* samples.

$$PDF_k^n(s) = \frac{1}{n\sigma} \sum_1^n W\left(\frac{s - s_k}{\sigma}\right) \tag{14}$$

From the result table, it is experimentally proved that for epilepsy identification in EEG signal, PNN gives the most accurate result by taking minimum amount of time.

3.2.3 Radial basis function neural network

RBF networks are also a type of feed-forward network, trained by using a supervised training algorithm. The main advantage of RBF network is, it has only one hidden layer. The RBF network, usually trains much faster than back-propagation networks. This kind of network is less susceptible to problems with non-stationary inputs because of the behaviour of radial basis function hidden units. The general formula for the output of RBF network [53] can be represented as follows, if we consider the Gaussian function as the basis function.

$$y(x) = \sum_{i=1}^{M} w_i e^{\left(\frac{-(\|x - c_i\|)^2}{2\sigma^2}\right)} \tag{15}$$

where x, $y(x)$, c_i, σ, and M denotes input, output, center, width, and number of basis function centered at c_i, similarly w_i denotes weights.

For this work, we have constructed a Radial Basis Function Network by taking into consideration of the Gaussian function as the basis function with a pre-fixing of randomized centres and widths.

3.3 Support vector machine (SVM)

SVM is the most widely used machine learning technique based pattern classification technique nowadays. It is based on statistical learning theory and was developed by Vapnik in the year 1995. The primary aim of this technique is to project nonlinear separable samples onto another higher dimensional space by using different types of kernel functions. In late years, kernel methods have received major attention, especially due to the increased popularity of Support Vector Machines [27]. Kernel functions play a significant role in SVM [28, 29] to bridge from linearity to nonlinearity. Least square SVM [30] is also an important SVM technique that can be applied for classification task [31]. Extreme learning Machine and Fuzzy SVM [32, 33, 34] and Genetic algorithm tuned expert model [32] can also be applied for the purpose of classification.

In this analytical work, we have evaluated three different types of kernel functions [35], i.e., Linear, Polynomial, and RBF kernel [36]. Linear kernel is the simplest kernel function available. Kernel algorithm using a linear kernel is often equivalent to their non-kernel counterparts [37]. From the result table it can be clearly understood that for a classification problem consisting of only sets A & E or D & E, is providing

100% accuracy. But it is not able to classify properly by considering sets A+D & E.

Polynomial kernel is a non-stationary kernel. This kernel function can be represented as given in equation:

$$K(x, y) = (\alpha x^T y + c)^d, \quad (16)$$

where α, c and d denotes slope, any constant, and degree of polynomial, respectively.

Somehow this kernel function [38, 39] is better as compared to linear kernel function. However, the RBF kernel function [40] has been proven as the best kernel function used for this application, which can classify different groups with 100% accuracy with a minimum time interval.

4 Ensemble of machine learning classifiers

From the empirical analysis we conclude that there some classifier models e.g., SVM and PNN outperforms all other techniques such as MLPNN, RNN, and RBFNN. Hence, to boost the poorly performer as compared to SVM and PNN, we have proposed a model for an ensemble based classifier that combines the above three techniques and improves the accuracy of classification for epileptic seizure detection. This proposed ensemble technique uses a weighted majority vote for classification. The main goal of an ensemble method is to combine the efficiency of several basic classifier models and build a learning algorithm to improve the robustness over a single classification technique. Here we have combined the classification results of MLPNN, RNN, and RBFNN and constructed an ensemble classifier. This classifier uses a weighted majority based vote for classification. Generally in majority based vote the class label of a sample is decided by the class label that is classified by maximum number of classifiers. Let there are two classes (1 and 2) and three classifiers (clf1, clf2, and clf3). Let for a sample clf1 and clf2 classifies to class 2 whereas clf3 classifies to class 1 then the ensemble of classifiers classifies the sample to class 2.

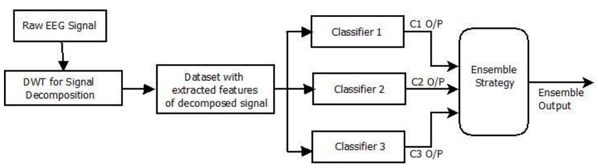

Figure 5: Proposed framework for ensemble based classifier for detection of epileptic seizure.

Figure 5 describes the architecture of ensemble of classifiers to detect epileptic seizure. It combines the output of three different classifiers such as classifier 1 (MLPNN), classifier 2 (RNN), and classifier 3 (RBFNN) based on a weighted majority voting mechanism. Weight parameter has been added to give different weightage to different classifiers based on their performance. For this we have collected the predicted class probabilities for each classifier and multiplied it with classifier weight and the average is taken. Based on this weighted average probabilities, the class label has been assigned. Here we have taken simple weighted majority technique where same weight is assigned to each class label which is 1/k, where k is the number of class labels.

5 Empirical study

This section gives an empirical study on different classification techniques based on machine learning approach for detection of epilepsy in EEG brain signal. Various experiments are done to validate this empirical study. The Machine Learning based classifiers are proved as the most efficient way for pattern recognition. It aids to design models that can learn from some previous experience (known as training) and further it can be able to recognize appropriate patterns for unknown samples (known as testing).

All experiments for this research work are performed using a powerful Java Framework known as Encog [54] developed by Jeff Heaton and his team. Currently we are using Encog 3.2 Java framework for all experimental result evaluation. This is the latest version and it supports almost all the features of machine learning techniques. Along with this framework, there are a lot of packages, classes and methods that have been defined to support the experimental evaluations. Java is the most potent and efficient language nowadays. The rightness of the experimental works can be verified easily using this language. There are almost nine different machine learning algorithms that have been implemented for EEG signal classification for epileptic seizure detection.

5.1 Environment and parameter setup

The Encog Java framework provides a vast circle of library classes, interfaces, and methods that can be utilized for designing different machine learning based classifier models. There are lists of parameters (as shown in Table 2) required to be set for smooth and accurate execution of models.

5.2 Performance measures and validation techniques

Here, we have hashed out about the performance of all machine learning based classifiers for classifying EEG signal. The different measures used for performance estimation are: Specificity (SPE), Sensitivity (SEN), Accuracy (ACC), and Time elapsed for execution of models. From the evaluation result given in Table 3, it is clear that MLPNN with resilient propagation is the most efficient training algorithm both in considerations of accuracy as well as the amount of time needed to execute the programs in all different setting such as A&E, D&E, and A+D & E. This MLPNN technique can be compared with other machine learning techniques. In this work all the experimental evaluations are validated using k- fold cross validation

Table 2: Lists of parameters for models execution.

Classification Techniques	Required Parameters and Values
MLPNN/BP	Activation Function - Sigmoid Learning Rate = 0.7 Momentum Coefficient = 0.8 Input Bias – Yes
MLPNN/RPROP	Activation Function - Sigmoid Learning Rate = NA Momentum Coefficient = NA Input Bias – Yes
MLPNN/MUR	Activation Function - Sigmoid Learning Rate = 0.001 Momentum Coefficient = NA Input Bias – Yes
SVM/Linear	Kernel Type – Linear Penalty Factor = 1.0
SVM/Polynomial	Kernel Type – Polynomial Penalty Factor = 1.0
SVM/RBF	Kernel Type – Radial Basis Function Penalty Factor = 1.0
PNN	Kernel Type – Gaussian Sigma low – 0.0001 (Smoothing Parameter) Sigma high – 10.0 (Smoothing Parameter) Number of Sigma - 10
RNN	Pattern Type – Elman Primary Training Type – Resilient Propagation Secondary Training Type – Simulated Annealing **Parameters for SA** Start Temperature – 10.0 Stop Temperature – 2.0 Number of Cycles - 100
RBFNN	Basis Function – Inverse Multiquadric Center& Spread Selection – Random Training Type– SVD (Singular Value Decomposition)

where value of k is taken as 10. So the total dataset has been divided into 10 folds. Each fold is constructed with samples having almost same number from each class labels. In each iteration one fold is considered for testing the classifier and rest of the folds are taken for training the classifier. This is a very efficient validation technique as it rules out all possibilities of misclassification and gives an accurate efficiency measure.

Table 4 shows a comparison of different kernel types used for classification using Support Vector Machine (SVM). It is the most powerful and efficient machine learning tool for designing classifier model. This table clearly shows a very good result for SVM with RBF kernel.

Table 5 defines a list of experiments led by studying different forms of Neural Network, such as Radial Basis Function Neural Network, Probabilistic Neural Network, and Recurrent Neural Network. It suggests that the effectiveness of using PNN for classification of EEG signal for detecting epileptic seizures is promising.

5.3 Comparative analysis

Table 6 gives a detail empirical analysis of the performance of different classification techniques based on machine learning approaches. As discussed above in this experimental evaluation we have used 10-fold cross validation to validate the results of classification.

Table 7 gives the result of experimental evaluation for the proposed ensemble technique and results for individual classification techniques. Figure 6 gives a graphical representation of comparison of different individual machine learning techniques with ensemble based classification technique. These experimental result shows there is a remarkable increase in the accuracy for case 3 (A+D & E) along with other two cases.

Table 3: Experimental evaluation result of MLPNN with different training algorithms.

Cases for Seizure Types	Multi-Layer Perceptron Neural Network with different Propagation Training Algorithms											
	Back-Propagation				Resilient-Propagation				Manhattan-Update Rule			
	SPE	SEN	ACC	TIME	SPE	SEN	ACC	TIME	SPE	SEN	ACC	TIME
Case1 (A,E)	100	90.09	94.5	16.52	99.009	100	99.5	2.846	97.29	77.77	85	7.541
Case2 (D,E)	100	83.33	90	22.22	99.009	100	99.5	2.547	55.68	78.78	60	7.181
Case3 (A+D,E)	100	86.95	92.5	23.12	95.85	85.98	92.33	14.79	93.78	82.24	89.66	14.85

Table 4: Experimental evaluation result of SVM with different kernel types.

Cases for Seizure Types	Support Vector Machine with different Kernel Types											
	Linear				Polynomial				RBF			
	SPE	SEN	ACC	TIME	SPE	SEN	ACC	TIME	SPE	SEN	ACC	TIME
Case1 (A,E)	100	100	100	2.127	100	100	100	2.101	100	100	100	2.002
Case2 (D,E)	100	100	100	1.904	100	100	100	1.902	100	100	100	2.021
Case3 (A+D,E)	90.67	76.63	85.66	11.61	100	99.009	99.66	7.24	100	100	100	2.511

Table 5: Experimental evaluation result of RBFNN, RNN, PNN with different training algorithms.

Cases for Seizure Types	Other Types of Neural Network											
	RBF Neural Network				Probabilistic Neural Network				Recurrent Neural Network			
	SPE	SEN	ACC	TIME	SPE	SEN	ACC	TIME	SPE	SEN	ACC	TIME
Case1 (A,E)	83.076	65.925	71.5	2.051	100	100	100	0.967	77.173	73.148	75	10.31
Case2 (D,E)	100	97.08	98.5	1.828	100	100	100	0.977	64.705	71.604	67.5	13.29
Case3 (A+D,E)	92.30	66.41	81	2.928	100	100	100	1.616	67.346	66.666	67.333	19.58

Table 6: Comparative analysis of different machine learning classification techniques.

Machine Learning Classification Technique	Case-1 (set A & E)		Case-2 (set D & E)		Case-3 (set A+D & E)	
	Overall Accuracy in %age	Approximate Time taken in seconds	Overall Accuracy in %age	Approximate Time taken in seconds	Overall Accuracy in %age	Approximate Time taken in seconds
MLPNN/BP	94.5	16.527	90	22.226	92.5	23.127
MLPNN/RP	99.5	2.846	99.5	2.547	92.33	14.798
MLPNN/MUR	85	7.541	60	7.181	89.66	14.85
SVM/Linear	100	2.127	100	1.904	85.66	11.61
SVM/Ploy	100	2.101	100	1.902	99.66	7.24
SVM/RBF	100	2.002	100	2.021	100	2.511
PNN	100	0.967	100	0.977	100	1.616
RNN	75	10.31	67.5	13.29	67.33	19.58
RBFNN	71.5	2.051	98.5	1.828	81	2.928

Table 7: Comparative analysis of different machine learning classification techniques with ensemble based classifier.

Machine Learning Classification Technique	Case-1 (set A & E)		Case-2 (set D & E)		Case-3 (set A+D & E)	
	Overall Accuracy in %age	Approximate Time taken in seconds	Overall Accuracy in %age	Approximate Time taken in seconds	Overall Accuracy in %age	Approximate Time taken in seconds
MLPNN/RP	99.5	2.846	99.5	2.547	92.33	14.798
RNN	75	10.31	67.5	13.29	67.33	19.58
RBFNN	71.5	2.051	98.5	1.828	81	2.928
ENSEBLE CLASSIFIER	99.5	3.745	99.5	3.876	98.3	5.475

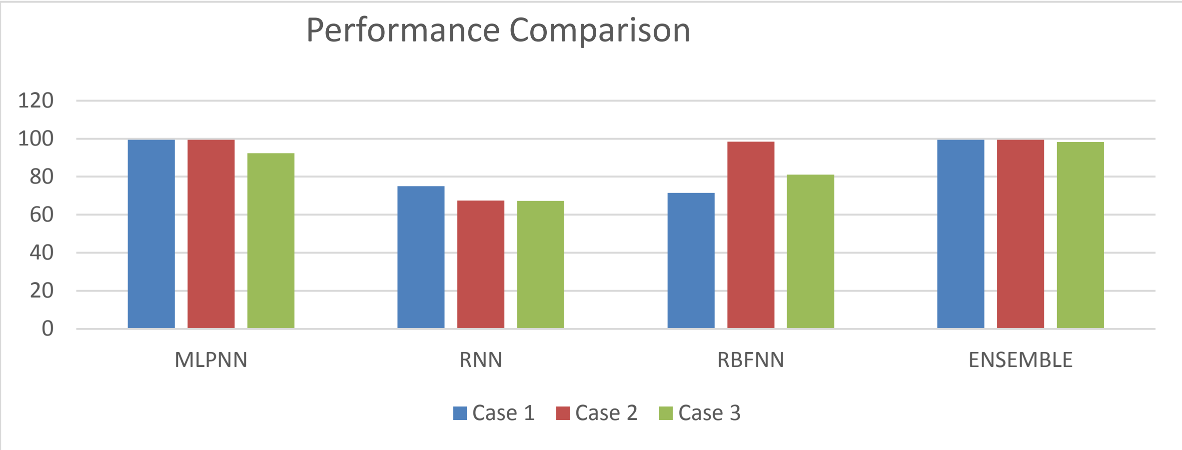

Figure 6: Performance comparison of different machine learning techniques with ensemble based classifier for detection of epileptic seizure.

6 Conclusions and future study

By classifying the EEG signals collected from different patients in different situations, detection of the epileptic seizure in EEG signal can be performed. Thus classification can be accomplished by using different machine learning techniques. In this work, the efficiency and functioning pattern of different machine learning techniques like MLPNN, RBFNN, RNN, PNN, and SVM for classification of EEG signal for epilepsy identification have been compared. Further, the tool MLPNN uses three training algorithms like BACKPROP, RPROP, and Manhattan Update Rule. Similarly, three kernels such as Linear, Polynomial, and RBF kernels are used in SVM. Hence, this comparative study clearly shows the differences in the efficiency of different machine algorithms with respect to the task of classification. Moreover, from the experimental study, it can be concluded that SVM is the most efficient and powerful machine learning technique for the purpose of classification of the EEG signal. Also, SVM with RBF kernel provides the utmost accuracy in all settings of the classification task. Besides this, PNN is a good contender for SVM for this specific application. But compared to SVM, PNN requires some extra overhead in setting the parameters. Also our proposed ensemble of classifiers based on weighted majority voting that combines the efforts of three different poorly performer classifiers such as MLPNN, RNN and RBFNN is enhancing the performance in different cases. Our continuous efforts in this area of research (both theoretical and experimental)is marching with lots of issues and will go ahead in future by considering the real cases with state-of-the-art meta-heuristic optimization techniques.

7 References

[1] Niedermeyer, E. and Lopesda Silva, F. (2005) *Electroencephalography: Basic Principles, Clinical Applications, and Related Field*s, 5th edition Lippincott Williams and Wilkins, London.

[2] Sanei, S. and Chambers, J. A. (2007) *EEG Signal Processing*, Wiley, New York.

[3] Lehnertz, K. (1999) 'Non-linear time series analysis of intracranial EEG recordings in patients with epilepsy--an overview', *International Journal of Psychophysiology*, Vol. 34, No. 1, pp.45-52.

[4] Alicata, F.M., Stefanini, C., Elia, M., Ferri, R., Del Gracco, S., and Musumeci, S.A. (1996) 'Chaotic behavior of EEG slow-wave activity during sleep', *Electroencephalography and Clinical Neurophysiology*, Vol. 99, No. 6, pp. 539–543.

[5] Acharya, U. R., Sree, S. V., Chuan Alvin, A. P., and Suri, J. S. (2012) 'Use of principal component analysis for automatic classification of epileptic EEG activities in wavelet framework', *Expert Systems with Applications*, Vol. 39, No. 10, pp.9072-9078.

[6] Acharya, U. R., Sree, S. V., Swapna, G., Joy Martis, R., and Suri, J. S. (2013) 'Automated EEG analysis of epilepsy: A review', *Knowledge Based System*, Vol. 45, pp. 147-165.

[7] EEG Data. [Online] *http://www.meb.unibonn.de/science/physik/eegdata.html*, 2001.

[8] Andrzejak, R. G., Lehnertz, K., Mormann, F., Rieke, C., David, P. and Elger, C. E.(2001) 'Indications of nonlinear deterministic and finite-dimensional structures in time series of brain electrical activity: Dependence on recording region and brain state', *Physical Review* E, Vol. 64, No. 6, pp. 1-6.

[9] Gandhi, T., Panigrahi, B. K., and Anand, S. (2011) 'A comparative study of wavelet families for EEG signal classification', *Neurocomputing*, Vol. 74, No. 17, pp. 3051–3057.

[10] Yong, L. and Shenxun, Z. (1998) 'The application of wavelet transformation in the analysis of EEG', *Chinese Journal of Biomedical Engineering*, pp. 333-338.

[11] Ocak, H. (2009) ‘Automatic detection of epileptic seizures in EEG using discrete wavelet transform and approximate entropy’, *Expert Systems with Applications*, Vol. 36, No.2, pp. 2027–2036.

[12] Adelia, H., Zhoub, Z., and Dadmehrc, N. (2003) ‘Analysis of EEG records in an epileptic patient using wavelet transform’, *Journal of Neuroscience Methods*, Vol. 123, No.1, pp. 69–87.

[13] Parvez, M. Z. and Paul, M. (2014) ‘Epileptic seizure detection by analyzing EEG signals using different transformation Techniques’, *Neurocomputing*, Vol. 145 (Part A), pp.190-200.

[14] Majumdar, K. (2011) ‘Human scalp EEG processing: various soft computing approaches, *Applied Soft Computing*, Vol. 11, No. 8, pp. 4433–4447.

[15] Teixeira, C. A., Direito, B., Bandarabadi, M., Quyen, M. L. V., Valderrama, M., Schelter, B., Schulze-Bonhage, A., Navarro, V., Sales, F., and Dourado, A. (2014) ‘Epileptic seizure predictors based on computational intelligence techniques: A comparative study with 278 patients’, *Computer Methods and Programs in Biomedicine,* Vol. 114, No. 3,pp. 324–336.

[16] Siuly and Li, Y. (2014) ‘A novel statistical algorithm for multiclass EEG signal classification’, *Engineering Applications of Artificial Intelligence,* Vol. 34, pp. 154–167.

[17] Jahankhani, P., Kodogiannis, V., and Revett, K. (2006) ‘EEG signal classification using wavelet feature extraction and neural networks’, Proc. of *IEEE John Vincent Atanasoff International Symposium on Modern Computing* (JVA 2006), pp.120-124.

[18] Guler, I.D. (2005) ‘Adaptive neuro-fuzzy inference system for classification of EEG signals using wavelet coefficients’, *Neuroscience Methods*, Vol. 148, pp. 113–121.

[19] Riedmiller, M. and Braun, H. (1993) ‘A direct adaptive method for Faster Back-propagation Learning: The RPROP Algorithm’, *Proc. of IEEE International Conference on Neural Networks,* Vol. 1, pp. 586 – 591.

[20] Mirowski, P., Madhavan, D., LeCun, Y., Kuzniecky, R. (2009) ‘Classification of patterns of EEG synchronization for seizure prediction’, *Clinical Neurophysiology*, Vol. 120, pp. 1927–1940.

[21] Subasi, A. and Ercelebi, E. (2005) ‘Classification of EEG signals using neural network and logistic regression’, *Computer Methods Programs in Biomedicine,* Vol. 78, No. 2,pp. 87–99.

[22] Subasi, A., Alkan, A., Kolukaya, E. and Kiymik, M. K. (2005) ‘Wavelet neural network classification of EEG signals by using AR model with MLE pre-processing’, *Neural Networks*, Vol. 18, No. 7, pp. 985–997.

[23] Ubeyli, E.D. (2009) ‘Combined neural network model employing wavelet coefficients for EEG signals classification’, *Digital Signal Processing*, Vol. 19, No. 2, pp. 297–308.

[24] Kalayci, T. and Ozdamar, O. (1995) ‘Wavelet pre-processing for automated neural network detection of EEG spikes’, *IEEE Engineering in Medicine and Biology Magazine*, Vol. 14, No. 2, pp. 160–166.

[25] Orhan, U., Hekim, M., and Ozer, M. (2011) ‘EEG signals classification using the K-means clustering and a multilayer perceptron neural network model’, *Expert Systems with Applications*, Vol. 38, No. 10, pp. 13475–13481.

[26] Pradhan, N., Sadasivan, P.K., and Arunodaya, G.R.(1996) ‘Detection of seizure activity in EEG by an artificial Neural Network: A Preliminary study’, *Computers and Biomedical Research*, Vol. 29, No. 4, pp. 303-313.

[27] Lima, C. A. M., Coelho, A. L. V., and Eisencraft, M. (2010) ‘Tackling EEG signal classification with least squares support vector machines: A sensitivity analysisstudy’, *Computers in Biology and Medicine*, Vol. 40, No. 8, pp. 705–714.

[28] Kumar, Y., Dewal, M.L. and Anand, R.S. (2014) ‘Epileptic seizure detection using DWT based fuzzy approximate entropy and support vector machine’, *Neurocomputing*, Vol.133, No. 7, pp. 271–279.

[29] Limaa, C. A.M. and Coelhob, A. L.V. (2011) ‘Kernel machines for epilepsy diagnosis via EEG signal classification: A comparative study’, *Artificial Intelligence in Medicine*, Vol. 53, No. 2, pp. 83–95.

[30] Siuly, Li, Y. and Wen, P. (2009) ‘Classification of EEG signals using sampling techniques and least square support vector machines’, *Rough Sets and Knowledge Technology*, Vol. 5589, pp. 375–382.

[31] Übeyli, E.D. (2010) ‘Least squares support vector machine employing model-based methods coefficients for analysis of EEG signals’, *Expert Systems with Applications*, Vol. 37, No. 1, pp. 233–239.

[32] Dhiman, R., Saini, J.S., and Priyanka (2014) ‘Genetic algorithms tuned expert model for detection of epileptic seizures from EEG signatures’, *Applied Soft Computing*, Vol. 19, pp. 8–17, 2014.

[33] Xu, Q, Zhou, H., Wang, Y., and Huang, J. (2009) ‘Fuzzy support vector machine for classification of EEG signals using wavelet-based features’, *Medical Engineering &Physics*, Vol. 31, No. 7, pp. 858–865.

[34] Yuan, Q, Zhou, W., Li, S., and Cai, D. (2011) ‘Epileptic EEG classification based on extreme learning machine and nonlinear features’, *Epilepsy Research*, Vol. 96, No. 1-2, pp. 29-38.

[35] Liu, Y., Zhou, W., Yuan, Q. and Chen, S. (2012) ‘Automatic Seizure Detection Using Wavelet Transform and SVM in Long-Term Intracranial EEG’, *IEEE Transactions on Neural Systems and*

Rehabilitation Engineering, Vol. 20, No. 6, pp. 749-755.

[36] Shoeb, A., Kharbouch, A., Soegaard, J., Schachter, S., and Guttag, J. (2011) 'A machine learning algorithm for detecting seizure termination in scalp EEG', *Epilepsy &Behavior*, Vol. 22, No. 1, pp. 36–43.

[37] Subasi, A. and Gursoy, M. I. (2010) 'EEG signal classification using PCA, ICA, LDA and support vector machines', *Expert Systems with Applications*, Vol. 37, No. 12, pp.8659–8666.

[38] Cristianini, N. and Taylor, J. S. (2001) *Support Vector and Kernel Machines*, Cambridge University Press, London.

[39] Taylor, J.S. and Cristianini, N. (2000) *Support Vector Machines and Other Kernel-Based Learning* Methods, Cambridge University Press, London.

[40] Vatankhaha, M., Asadpourb, V., and FazelRezaic, R. (2013) 'Perceptual pain classification using ANFIS adapted RBF kernel support vector machine for therapeutic usage', *Applied Soft Computing*, Vol. 13, No. 5, pp. 2537–2546.

[41] Pineda, F.J. (1987) 'Generalization of back-propagation to recurrent neural networks', *Physical Review Letters*, Vol.59, No. 19, pp. 2229–2232.

[42] Adeli, H. and Panakkat, A. (2009) 'A probabilistic neural network for earthquake magnitude prediction', *Neural Networks*, Vol. 22, No. 7, pp. 1018-1024.

[43] Ghosh-Dastidar, S., Adeli, H., and Dadmehr, N. (2008) 'Principal Component Analysis-Enhanced Cosine Radial Basis Function Neural Network for Robust Epilepsy and Seizure Detection', *IEEE Transactions on Biomedical Engineering*, Vol. 55, No. 2, pp.512-518.

[44] Guler, N. F., Ubeyli, E. D., and Guler, I. (2005) 'Recurrent neural networks employing Lyapunov exponents for EEG signals classification', *Expert Systems with Applications*, Vol. 29, No. 3, pp. 506–514.

[45] Petrosian, A., Prokhorov, D., Homan, R., Dascheiff, R., and Wunsch, D. (2000) 'Recurrent neural network based prediction of epileptic seizures in intra- and extracranial EEG', *Neurocomputing*, Vol. 30, No. 1-4, pp. 201–218.

[46] Petrosian, A., Prokhorov, D.V., Lajara-Nanson, W., and Schiffer, R. B. (2001) 'Recurrent neural network-based approach for early recognition of Alzheimer's disease in EEG', *Clinical Neurophysiology*, Vol. 112, No. 8, pp.1378–1387.

[47] Übeyli, E. D. (2009) 'Analysis of EEG signals by implementing Eigen vector methods/recurrent neural networks', *Digital Signal Processing*, Vol.19, No. 1, pp.134–143.

[48] Derya, E. (2010) 'Recurrent neural networks employing Lyapunov exponents for analysis of ECG signals', *Expert Systems with Applications*, Vol. 37, No. 2, pp. 1192–1199.

[49] Saad, E.W., Prokhorov, D.V. and Wunsch, D.C. (1998) 'Comparative study of stock trend prediction using time delay, recurrent and probabilistic neural networks', *IEEE Transaction on Neural Networks*,Vol.9, No. 6, pp. 1456–1470.

[50] Gupta, L., McAvoy, M., and Phegley, J. (2010) 'Classification of temporal sequences via prediction using the simple recurrent neural network', *Pattern Recognition*, Vol. 33,No. 10, pp. 1759–1770.

[51] Gupta, L. and McAvoy, M. (2000) 'Investigating the prediction capabilities of the simple recurrent neural network on real temporal sequences', *Pattern Recognition*, Vol. 33, No. 12, pp. 2075–2081.

[52] Derya, E. (2010) 'Lyapunov exponents/probabilistic neural networks for analysis of EEG signals', *Expert Systems with Applications*, Vol. 37, No. 2, pp. 985–992.

[53] Saastamoinen, A., Pietila, T., Varri, A., Lehtokangas, M. and Saarinen, J. (1998) Waveform detection with RBF network application to automated EEG analysis', *Neurocomputing*, Vol. 20, No. 1-3, pp. 1–13.

[54] Heaton, J. (2011) *Programming Neural Networks with Encog 3 in Java*, 2nd Edition, Heaton Research.