A Nearest Neighbor Based Approach for Classifying Epileptiform EEG Using NonLinear DWT Features

Ashwini V R Holla
holla_ashwini@yahoo.co.uk

Aparna P
p.aparnadinesh@gmail.com

Department of Electronics and Communication Engineering
National Institute of Technology Karnataka, Surathkal
Mangalore, India

***Abstract*—Epilepsy is a pathological condition characterized by spontaneous, unforeseeable occurrence of seizures, during which the perception or behaviour of a person is altered, if not disturbed. In prediction of occurance of seizures, better classification accuracies have been reported with the use of non linear features and hence they have been estimated from wavelet transformed Electro Encephalo Graph (EEG) data and used to train *k* Nearest Neighbour (kNN) classifier to classify the EEG into *normal, background* and *epileptic* classes. Very good accuracy performance of nearly 100% has been reported from the current work.**

Key Words: Electro Encephalo Graph (EEG); k Nearest Neighbour (kNN); non linear features ; epileptic seizure;

I. Introduction

A scheme for predicting the occurrence of seizures of epilepsy patients would help not only the patients from a potentially risky situation but also observers or onlookers to take precautions regarding patient safety. Most of the studies documenting methods for epilepsy detection have focused on better accuracies in classifying seizure or the ictal phase from the rest of the EEG recording rather than tapping characteristic features extracted from the EEG that are predictive of impending seizure(s). Thus the study was mainly geared for segregating a background epileptic state or the aura phase indicative of impending seizures from the EEG with as high accuracy as possible.

Proper diagnosis for epilepsy via gold standards necessitate the use of expensive imaging techniques as functional Magnetic Resonance Imaging (fMRI) to pinpoint the focii of epilepsy. With advances in computing speeds and power, it is now possible to not only judge patient condition objectively but also automate diagnostical decisions. The authors propose a scheme where in a classifier discriminates based on the features estimated from the detail coefficients of the wavelet decomposition of the data.

Epileptiform EEG

More often than not time series analysis of the signal is relied upon which fail in detecting significant characteristics that may be recorded in a different domain. The need for automation in EEG monitoring for seizure prediction and detection is imperative to manage the condition and to mitigate ambiguity brought in by human errors and observations, however skilled or otherwise [1]. It may also be noted that not all epilepsies present epileptiform EEGs, hence the scope for human error increases.

Distinguishing seizure signals from common artifacts is not very difcult as they are prominent spiky but repetitive in nature, whereas most other artifacts are transients or noise-like in shape. The ictal wave patterns, appear with the onset of epilepsy. Spikes have a high correlation with seizure occurrence and are usually of 10 - 80 ms duration. They are often followed by slow waves usually occuring at 3 Hz or lesser frequencies. The background epileptic state encompasses pre-ictal or interictal phases that occur prior to the onset of siezures and is often called as the *aura*indicative of impending seizure. Fig. 1 shows plots of Normal, Pre-Ictal and Ictal EEG of 5 seconds duration.

Schemes for epilepsy detection are briefly discussed:

Spectral analysis: Peaks in the Power Spectrum at specific frequencies may be used to identify epileptic seizures. Estimating the power spectrum may be carried out nonparametrically using the Fourier transform of the estimate of the autocorrelation or by the parametric approach that uses a model for the same[2].

Local variance: The signal is tessellated into rectangular windows and variances in each segments are compared with a common threshold, if greater than the threshold, the segment is has a seizure record else, is normal [3]. Thresholds may be computed adaptively rather than being a fixed entity[4].

Transform domain Analysis: Although any orthogonal transform can be used, wavelets are gaining faster and newer grounds in recent times with good results reported. In [5] EEG was analysed with 5 level decomposition using Daubechies 4 wavelet filter and classified using neural networks with the energy of details and approximations being the input features. In [4], variances and standard deviation of wavelet detail coefficients have been used for epilepsy detection.

978-1-4673-2014-6/12/$31.00 ©2012 IEEE

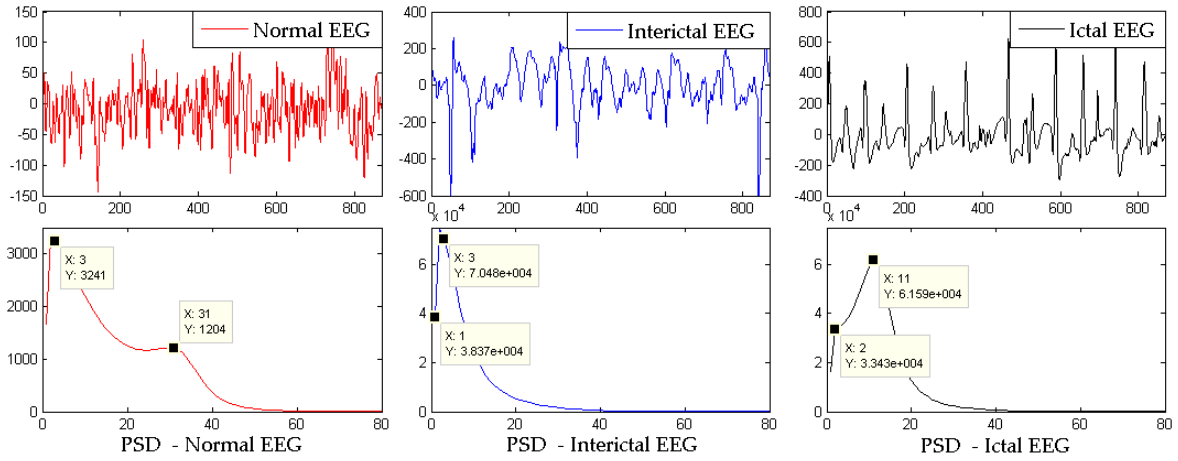

Fig. 1: Typical Plots of Normal Interictal & Ictal EEG taken for a duration of 5 s along with their PSD's in second row.

Nonlinear measures: Nonlinear measures like Correlation Dimensions (CD), Largest Lyapunov exponents (LLE), entropies, help to understand the EEG dynamics and underlying chaos in the brain [6]. In [7], CD was used to characterize the background epileptic EEG for seizure prediction. On comparison non linear measures fare better than spectral analysis. Seizure detection performance of various entropy measures tested in [8] and entropy values computed for the epileptic EEG were found to be lower compared to the values computed for the normal EEG. It has also been reported that EEG data during seizure activity has significant non linearity despite which, it is more predictable than seizure free intervals that resemble Gaussian linear stochastic processes, as expressed in [9].

II. Proposed Method

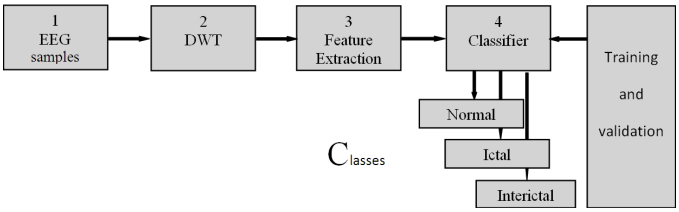

Fig. 2: The Proposed method

The approach is as illustrated in Fig. 2. The EEG data used in this research have been made available at the site [10] with details at [7].

A. Discrete Wavelet transform (DWT)

Multiresolution analysis of a signal through DWT enables one to analyse data with different levels of detail at different scales of decomposition. Transforming the signal into the DWT domain needs the data to be sampled first. Choosing a higher sampling frequency increases the computational complexity, given by Mallats algorithm [11] : $O(2^d nk)$ where : d - depth of decomposition, n - number points in the data and k - number of filter coefficients. It is interesting to note that higher number of decomposition levels and higher sampling frequencies do not translate to higher accuracies.

B. Feature Extraction

The choice of entropies for feature extraction stems from the fact that epileptic EEG exhibit a high rate of periodicity, decreasing randomness hence, the measure of information during epilepsy. The following are the 5 entropy estimators used. **Approximate Entropy Estimator:** $ApEn(m, r, N)$ of a series is considered given: Run length ***m***,Tolerance window ***r***, Number of sample points ***N*** $ApEn$ measures the log likelihood that series of patterns close (within tolerance) for a given number of consequent observations are close on incremental comparisons that follow[12]:

$$ApEn(m, r, N) = \phi^m(r) - \phi^{m+1}(r) \tag{1}$$

$$\phi^m(r) = \frac{1}{N-m+1} \sum_{i=1}^{N-m+1} ln(C_r^m(i))$$

where:

$$C_r^m(i) = \frac{N^m(i)}{N-m+1}$$

$C_r^m(i)$ values measure, within a tolerance r, the frequency of occurance of patterns similar to a given one of window length m and is called 'correlation integral'

Sample entropy Estimator: The $SampEn(m, r, N)$ is the negative logarithm of the conditional probability that two sequences with similar ***m*** points will remain similar at the next consecutive point. The advantage here is that more often than not, irrespective of record length, $SampEn$ results are relatively consistent where $ApEn$ results are not [13],[14].

$$SampEn(m, r, N) = -ln\Big[\frac{A^m(r)}{B^m(r)}\Big] \tag{2}$$

$$A^m(r) = \frac{1}{N-m} \sum_{i=1}^{N-m} A_i^m(r); B^m(r) = \frac{1}{N-m} \sum_{i=1}^{N-m} B_i^m(r)$$

Renyi Entropy Estimator: Renyi entropy of the order α given $\alpha \geq 0$ and $\alpha \neq 1$, for a discrete random variable $X_N = x_1, x_2....x_N$ with p_i the probabilty of occurance of the event $X = x_i$ is given modified from [13] and [8] as:

$$H\alpha(X_N) = \frac{1}{1-\alpha} \sum_{i=1}^{N} log\{p_i^{\alpha}\} \tag{3}$$

Higher Order Spectrum (HOS) Entropy estimators: These are normalised entropy estimators from poly spectra and are representations of higher-order moments or cumulants of a signal. The ***bispectrum*** of a signal is the Fourier transform of the third-order correlation of the signal. It can be estimated using the averaged biperiodogram given by [16]

$$B(f_1, f_2) = E[X(f_1)X(f_2)X^*(f_1 + f_2)] \tag{4}$$

$X(f)$is the Fourier transform of the signal $x(nT)$ and $*$ denotes complex conjugation and $E[.]$ denotes the expectation operation.

The Fourier transform of a real-valued signal shows conjugate symmetry, and the power spectrum is redundant in the negative frequency region. Likewise the ***bispectrum*** being a product of three Fourier coefficients, exhibits symmetry and is computed in the non-redundant region, Ω as indicated in Fig.3. Formulae

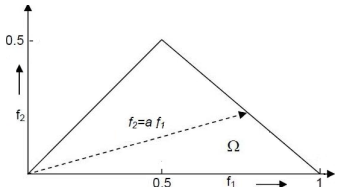

Fig. 3: Nonredundant region in PSD of the bispectrum

for these bispectral entropy estimators taken from [16] are given:

Normalized bispectral entropy 1, P_1:

$$P_1 = \sum_n p_i log p_i; p_i = \frac{|B(f_1, f_2)|}{\sum_{\Omega} |B(f_1, f_2)|} \tag{5}$$

Similarly Normalised Bispectral entropy 2, P_2 is:

$$P_2 = \sum_n p_n log p_n; p_n = \frac{|B(f_1, f_2)|^2}{\sum_{\Omega} |B(f_1, f_2)|^2} \tag{6}$$

HOS analysis helps detect non-linearity and phase relationships between harmonic components and characterises regularity of physiological signals much better than its peers [15].

C. Statistical Significance test

Before proceeding to classify, one needs to determine the statistical significance of the features obtained. A *Student's t test* or *Analysis of Variance* (ANOVA) test or any other hypothesis test that measures how close the distributions of the three classes lie can be used. The ***p*** value provides evidence in support of the null hypothesis and is the probability of obtaining the study results ***p(F)*** if the null hypothesis is true. Small ***p*** values indicate disjoint groups with little overlap in distributions indicating better statistical significance. Given in Section IV is Table. I that contains ***p*** values of the five entropies discussed.

Basic *k*NN Algorithm

Input:
D - Set of training objects,
z - test object- a vector of attribute values.
L - Set of classes used to label the objects.
Output : $C_z \in \mathbf{L}$ - the class of ***z***.
for each object $\boldsymbol{y} \in \boldsymbol{D}$
do
| Compute d(***z, y***) - the distance between ***z*** and ***y***
end
Select N $\subseteq$ D, the set (neighborhood) of ***k*** closest training objects for ***z***;
$C_z = argmax_{v \in L} \sum_{y \in N} I(v = class(c_y))$;
I(·) - an indicator function; returns the value 1 if its argument is true else 0.

D. Classifier

A classifier is a routine that takes in a set of data with assigned labels called classes and tries to group another new similar data set under the given classes based on some decision rule that considers some key features of the data sets. For the current work, *k Nearest Neighbor* (***k*NN**) algorithm has been chosen. It finds a cluster of ***k*** datapoints in the training vector closest to the test object and classifies test object based on class of the majority of the neighbors. Ties are broken in a manner specified, for by taking the class of the most frequent class in the training set. Given inside the box is the algorithm adapted from [17].

III. Data

The test subjects were 5 healthy and 5 epileptic patients diagnosed with *temporal lobe epilepsy*. While 200 normal EEG data samples were obtained using an Internationally standardized 10 20 surface electrode placement scheme with from 5 healthy volunteers in two states: 100 data samples with relaxed awake state with eyes open and another 100 with eyes-closed. The 100 ictal samples were recorded *intracranially*, with electrodes placed on the correct *epileptogenic* zone, during epileptic seizures of 5 epileptic patients, while the 200 pre-ictal EEG readings taken intracranially were from the same 5 patients when they exhibited no seizure activities with electrodes placed on the *epileptogenic* site and on a site, its polar oppposite (100 each). Each of the 23.6 s duration, single channel artifact free recordings were sampled at 173.61 Hz, digitized with 12-bit Analog to Digital Converter and encoded in ASCII.

IV. Results

Initially the 52 wavelets from the following 7 families - Haar, Meyer, Daubechies(2-10), Coiflets(1-5), Symlets(2-7), Biorthogonal and Reverse Biorthogonal(1.1, 1.3, 1.5, 2.2, 2.4, 2.6, 2.8,3.1, 3.3, 3.5, 3.7, 3.9, 4.4, 5.5, and 6.8), were used

to decompose the data to six levels, but after the statistical significance test, only first level of detail coefficients were selected. The data was segmented with a 5 second window to improve upon the stationarity. $ApEn$ values were estimated with m value 5 and r assigned 0.2. The same values were used for estimating $SampEn$ while α was assigned a value 2 for *Renyi*'s entropy .

A. Test vector generation and Cross Validation

Segmenting each data file into sets of 5 seconds provided for a total of 2400 files from three classes of which 1680 were used for training and 720 for testing using a three fold cross validation. The scheme was tested on Matlab®2006a with 1 nearest neighbour and *Euclidean distance metric* and nearest neighbour rule to mitigate contentions. From the results, it was seen that *biorthogonal 2.4* filter yielded the best results of which the Mean, standard deviation and ***p*** value for all of the 5 features are tabulated in Table.I.

Performance Metrics The following performance metrics are usually selected to assess classifiers:

Precision = $TP/(TP+FP)$Percentage of correct $+$ve predictions.

Recall/Sensitivity = $TP/(TP+FN)$ Percentage of $+$ve labeled instances predicted as $+$ve.

Specificity = $TN/(TN+FP)$ Percentage of $-$ve labeled instances predicted as $-$ve.

Accuracy = $(TP+TN)/(TP+TN+FP+FN)$ Percentage of correct predictions.

True positives (TP): Seizures identified by the classifier and EEG experts.

False positives (FP): Seizures identified by the classifier but not experts.

False negatives (FN):Seizures missed by the classifier system.

True Negatives (TN): Non Seizures identified by both parties.

Table.II gives the performance metrics, along with average time taken to classify the data using 3 level cross validation on an Intel® 2GHz dual core processor running on MS Windows XP®.

B. Discussion

In light of the ensuing discussion about the features that have so far been reported to have given the best results , validity of the results obtained from the current work can be established as being superior. In TableIII 5 papers documenting ***3 class*** EEG classification schemes using the same database except for the last paper [21] which uses binary classification are summarily compared:

1) In [18], the data is decomposed to 5 levels with a Daubechies 25 order filter. and classified with a K means classifier. The accuracy got by constructing the filter that requires 50 coefficients is around 97% .
2) [19] has statistical features of the detail coefficents of the Daubechies 2 DWT at 4th level decomposition along with LLE being used to train *Modified Mixture of Experts* (MME). 6 expert networks (*multilayer perceptron neural networks*) constituted the MME.
3) [20] discusses how after 4 level decomposition using a Daubechies 4 filter, the statistical features as Standard Deviation and Non Linear Chaotic parameters as CD and LLE can be used to train the classifiers. The coefficients of the 4 bands are fed to 4 different classifiers in different ways. *Unsupervised k -means clustering* and *statistical discriminant analysis, radial basis function neural network* and *LevenbergMarquardt Back Propagation Neural Network* (LMBPNN) classifiers are used. The computation complexity and time taken do not justify the accuracy a maximum possible of 96% with just 20 epochs tested.
4) In the study described in [16] *Gaussian Mixture Models* (GMM) and *Support Vector Machines* (SVM) classifiers are trained with the bispectrum phase invariant features mean *bispectrum magnitude, bispectrum phase entropy,* and normalized *bispectrum entropies* all of them being features derived from the bispectrum, The average accuracy being 93.11 % and 92.70 %, respectively. The pre-ictal EEG was classified with 89.67% and 84.00% accuracy by GMM and SVM classifiers respectively.
5) [21] Uses $ApEn$ alone as a feature that trains a *Probailistic Neural Network* (PNN) classifier. By far the best classification accuracy of 99.6% can be obtained, but with zero tolerance and with N=512 making the data segment 3 second long. Approximate entropy results are unreliable for short data and also only 77.42% overall accuracy is reported for a binary classification (Considering all values of N). Though the main feature of the scheme is its simplicity, the results cannot be used for practical purposes as the features selected are likely to give unreliable results for small sequences.

There are papers besides these that have reported accuracies as good as the authors have presented but consider binary classification and use time segments much smaller than 1 s. The three factors for the good results have been the use of non linear features and the biorthogonality of the wavelet basis along with the use of smaller sequence of data. While segmenting the data improved accuracy, taking too small time frames may again lead to precarious results.

V. Conclusion and Future work

In this paper is presented a scheme of using k Nearest Neighbour classifier trained with non linear features of DWT of EEG samples to detect and predict the onset of epileptic seizures and has been objectively evaluated using four metrics.

The outstanding results may be attributed to the orthogonality property of the wavelet under consideration namely *biorthogonal 2.4*, along with the use of the non linear features discussed, further windowing of the data has contributed to boosted performance.

All features have been taken from just the first level of *biorthogonal 2.4* wavelet thereby decreasing the complexity as well as constraints on the sampling frequency [11]. This reduction in complexity along with the use of features simple in implementation, promise the potential use of the system in clinical application after due modifications.

TABLE I: '***p***' values with Mean, Standard deviation(SD) and F Statistic for the features taken from *biorthogonal 2.4* wavelet

	Renyi		**ApEn**		**SampEn**		P_1		P_2	
♠	**Mean**	**±SD**	**Mean**	**±SD**	**Mean**	**±SD**	**Mean**	**±SD**	**Mean**	**±SD**
Ictal	7.526	1.7336	1.804	0.226	1.394	0.416	0.916	0.035	0.729	0.123
Background Epileptic	5.225	2.674	2.0259	0.209	2.077	0.344	0.954	0.009	0.843	0.038
Normal	6.209	2.476	2.103	0.079	2.160	0.124	0.928	0.029	0.737	0.122
'p' value	**0**		**0**		**0**		**0**		**0**	
F statistic		**196.38**		**573.97**		**1384.66**		**445.03**		**309.19**

TABLE II: Performance Metrics along with time for computation of bior2.4 features using **kNN** classifier

Accuracy with non DWT features	**Accuracy with DWT features**	**Precision**	**Sensitivity**	**Specificity**	**Avg CPU time in seconds**
91.5	**99.6**	**99.9**	**99.4**	**99.9**	**0.54**

TABLE III: Table of Performance metrics comparing recent papers and the method proposed

Features used to train Classifiers or method used to classify	**Sensitivity**	**Specificity**	**Accuracy**	**Author papers**
Variance of 5^{th} level Daub 25 coeffecients - K Means classifier	96.1	97.67	99	Suparerk J in [18]
Statistical features of 4^{th}level Daub2 coeffecients and LLE - MME	97.5Preictal 98 Ictal	98 Normal	98.5 Normal	Ubeyli E in [19]
Mixed Band Wavelet Chaos Neural Network		96.67		Samonwoy.G, Adeli H, Nahid.D in [20]
HOS features - GMM and SVM	96.67(GMM) 95.67(SVM)	93.11(GMM) 92.67(SVM)	92(GMM) 96.67(SVM)	Chua K, U Rajendra Acharya, Eric C, Lim C.M, Toshiyo Tamura in [16]
ApEn - PNN		99.6		Srinivasan. V, Eswaran. C, Sriram N in [21]
Non linearfeatures of DWT fed to kNN classifier	**99.4**	**99.9**	**99.6**	**Current work**

ACKNOWLEDGEMENT

The Authors wish to acknowledge the guidance of Prof U Rajendra Acharya especially in developing HOS entropies. The authors would also like to acknowledge the work of Prof Andrzejak, Bonn University, Neurology Department, Centre for Epilepsy in recording the data and making it available for the public.

REFERENCES

[1] K. Radhakrishnan, B. Santoshkumar, A.Venugopal, "Prevalence of benign epileptiform variants observed in an EEG laboratory from South India", *J. Clin Neurophysiol* 1999, 110: pp.280-285.

[2] Oliver Faust, U. Rajendra Acharya, Lim C.M, Bernhard H.Sputh, "Automatic Identification of Epileptic and Background EEG Signals Using Frequency Domain Parameters", *Int. J. Neural Syst*, 2010 vol. 20, no. 2, pp. 195-176.

[3] Tarassenko L, Khan Y. U, Holt M. R. G, "Identification of inter-ictal spikes in the EEG using neural network analysis," *Inst. Elect.Eng.Proc. Sci. Meas. Technol*, Nov.1998 vol. 145, no. 6, pp. 270278.

[4] Indiradevi K.P, Elizabeth Elias, Sathidevi P.S, Nayak S. Dinesh, Radhakrishnan K. "A multi-levelwavelet approach for automatic detection of epileptic spikes in the electroencephalogram" *Computers in Biology and Medicine* Vol. 38, 7, 2008, pp.805 - 816 .

[5] Subasi A, "Application of adaptive neuro-fuzzy inference system for epileptic seizure detection using wavelet feature extraction", *Computers in Biology and Medicine*, 2007, 37(2), 227244 .

[6] Kannathal N, U. Rajendra Acharya, Lim C.M, Sadasivan P.K, "Characterization of EEG - A comparative study ",*J.Computer Methods and Programs in Biomedicine* ,Oct 2005 , Vol. 80, Issue 1, pp 17-23.

[7] Andrzejak R.G, "Indications of nonlinear deterministic and finite dimensional structures in time series of brain electrical activity: Dependence on region and brain state", *Physical Review E*, 2001 vol. 64(6).

[8] Kannathal N, Lim C.M, U. Rajendra Acharya, SadasivanP.K, "Entropies for detection of epilepsy in EEG ",*J.Computer Methods and Programs in Biomedicine*, Dec. 2005, Vol. 80, Issue 3, pp. 187-194.

[9] Hasan Ocak, "Automatic detection of epileptic seizures in EEG using discrete wavelet transform and approximate entropy", *J Expert Systems with Applications* 36, 2009, 20272036, Elsiever Publications.

[10] The data from Bonn University is available at http: www.meb.uni-bonn.de/epileptologie/science/physik/eegdata.html.

[11] Stéphane Mallat, *" A Wavelet Tour of Signal Processing"*, San Diego: Academic Press, 1999.

[12] Pincus, S. M., "Approximate entropy as a measure of system complexity", *Proc. Natl. Acad. Sci.USA.*, vol. 88, pp. 2297-2301.

[13] Costa M, Goldberger A.L, Peng C.K, "Multiscale entropy analysis of biological signals", *Phys Rev E* 2005,71:021906.

[14] Richman, Joshua S, Moorman, J. Randall, "Physiological time-series analysis using approximate entropy and sample entropy", *Am J Physiol Heart Circ Physiol*, 2000, vol. 278, pp. H2039–2049.

[15] David R. Brillinger, "Some basic aspects and uses of higher-order spectra", *Signal Processing*, 1994 vol. 36, 3, pp. 239-249.

[16] U. Rajendra Acharya, Eric Chern-P.C, Chua K.C, Lim C.M, Toshiyo Tamura, "Analysis and Automatic Identification of Sleep Stages Using Higher Order Spectra", *Int. J. Neural Syst*, 2010, vol. 20, no. 6, pp 509-521.

[17] Wu X, Kumar V, Ross Quinlan, Ghosh J, Yang Q, Motoda H, McLachlan G.J,Ng A, Liu B, Yu P.S and others, "Top 10 algorithms in data mining " Springer,*J. Know. Inf Sys*, Vol 14,1,p(1-37),2008.

[18] Suparerk Janjarasjitt, "Classification of the Epileptic EEGs Using the Wavelet-Based Scale Variance Feature",*J. Applied Biomedical Engineering*, 2010, Vol.3,1, pp 19-25.

[19] Ubeyli, E.D, "Modified Mixture of Experts for Analysis of EEG Signals," *Engineering in Medicine and Biology Society*, EMBS 2007. 29th Annual International Conference of the IEEE , pp.1546-1549, 22-26 Aug. 2007

[20] Samanwoy Ghosh-Dastidar, Hojjat Adeli, Dadmehr Nahid, "Mixed-Band Wavelet-Chaos-Neural Network,Methodology for Epilepsy and Epileptic Seizure Detection", *IEEE Transactions on Biomedical Engineering, 2007*, vol. 54, 9, pp.1545-1551.

[21] Srinivasan V, Eswaran C, Sriram N, "Approximate Entropy-Based Epileptic EEG Detection Using Artificial Neural Networks",*IEEE Transactions on Information Technology In Biomedicine*, May 2007, Vol. 11, 3, pp.288-295.