Intelligent Classifiers of EEG Signals for Epilepsy Detection

AYMAN M. MANSOUR
Department of Communication, Electronics and Computer Engineering,
Tafila Technical University, Tafila, 66110, JORDAN

MOHAMMAD A OBEIDAT
Department of Power and Mechatronics Engineering,
Tafila Technical University, Tafila, 66110, JORDAN

MURAD AL-AQTASH
Department of Communication, Electronics and Computer Engineering,
Tafila Technical University, Tafila, 66110, JORDAN

Abstract: - In this paper a novel methodology is developed to classify the time series EEG signals, and to apply that methodology to the analysis of physiological signals recorded from epileptic patients for seizure analysis depending on EEG signal. The developed system uses data mining techniques as classification tool based on standard deviation, entropy, and power spectral density resulted of EEG signal components resulted from discrete wavelet transform. The work has been tested on real patient EEG signals using Weka software. This approach offers a good treatment of seizure detection.

Key-Words: - Epilepsy, Standard Deviation, Entropy, Power Spectrum density, Discrete Time Wavelet Transform.

1 Introduction

The electroencephalogram (EEG) signals have long been recorded and studied as potentially of the electric activity of the brain. It has stood the test of time because it monitors physiological functions as these changes with time and is reactive to many factors. EEG reflect brain electrical activity with millisecond temporal resolution, and are the most direct correlate of on-line brain processing obtainable non-invasively.

The electrical activity of active nerve cells in the brain produces currents spreading through the head. These currents also reach the scalp surface, and resulting voltage differences on the scalp can be recorded as the electroencephalogram (EEG). The (EEG) signal contains–among others-four main spectral components. The most important one is the α wave, having the frequency between 8-12 Hz . Higher frequency components (over 12 Hz) are denoted by β, these waves have lower magnitudes. Lower frequency components δ (under 4 Hz) and θ(between 4-8 Hz) are also present, both having low magnitudes in normal cases [1,2].

The clinical interests in (EEG) are; for example, sleep pattern analysis, cognitive tasks registration, seizure and epilepsy detection, and other states of the brain, both normal and patho-physiological. Epilepsy is the second most prevalent neurological disorder in humans after stroke. It is characterized by recurring seizures in which abnormal electrical activity in the brain causes altered perception or behavior. Well-known causes of epilepsy may include: genetic disorders, traumatic brain injury, metabolic disturbances, alcohol or drug abuse, brain tumor, stroke, and infection.

A seizure is a disturbance characterized by changes in neuronal electrochemical activity that results in abnormal synchronous discharges in a large cell population, giving rise to clinical symptoms and signs. These are some of the seizure patterns that one sees on EEG: 1) Absence seizures have the characteristic 3hz spike and wave discharges. 2)Juvenile myoclonic epilepsy has a 4Hz spike and wave discharge pattern 3)Infantile spasms show the ominous pattern of hypsarrhythmia 4) Lennox Gastaut shows 1-2 Hz spike and wave discharges 5) Focal spikes and spike wave complexes

In addition to the characteristic electrographic bursts of abnormal activity that are recorded when epileptic patients experience a seizure (ictal episode), the electroencephalogram (EEG) of epileptics will normally display isolated sharp transients or "spikes" in some locations of the brain. These interictal spikes are a complementary source of information in the diagnosis and localization of epilepsy [3-8].

The EEG of epileptics will normally display isolated sharp transients or "spikes" in some locations of the brain. These spikes are a main source of information in the diagnosis and localization of epilepsy. Fig. 1 shows the EEG signal for normal case. Fig.2 shows epileptic EEG signal.

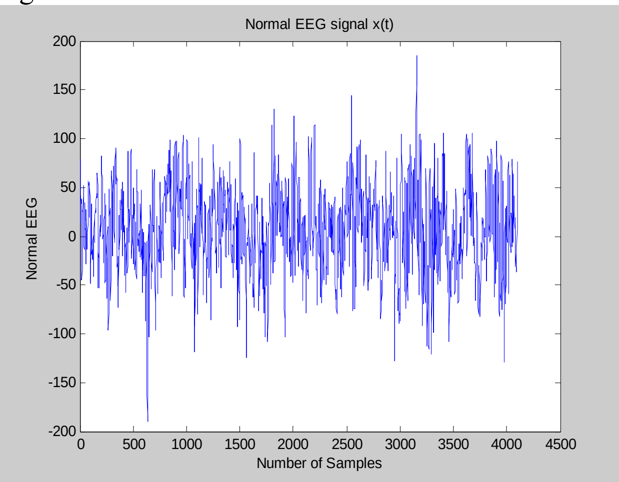

Fig. 1. Normal EEG Signal

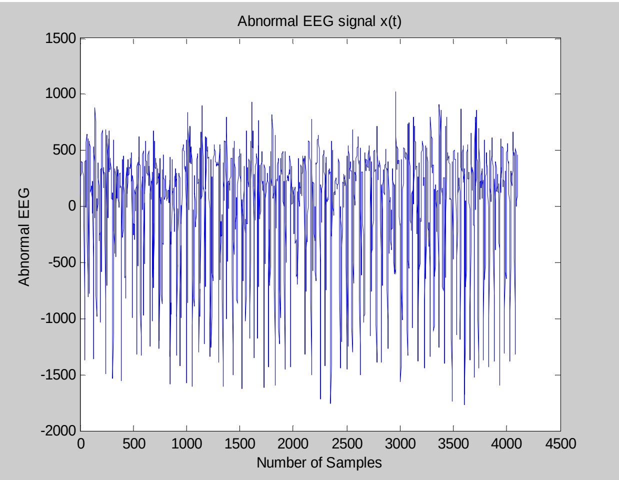

Fig. 2. Epileptic EEG Signal

2 Methodology

EEG signal classification process involves four phases: Data gathering phase, Data preprocessing phase, the learning phase and the recognition phase. In data gathering, the training and test set will be obtained from EEG signal databases. The second phase is to pre-process the signals, including sampling, Whitening Centering, Dimensionality Reduction. In the learning phase, the target is to build a model. The last step is using remainder of pre-processed data to test the model. A test set is used to determine the accuracy of the model. Usually, the given signals is divided into training and test sets, with training set used to build the model and test set used to validate it (Fig.3).

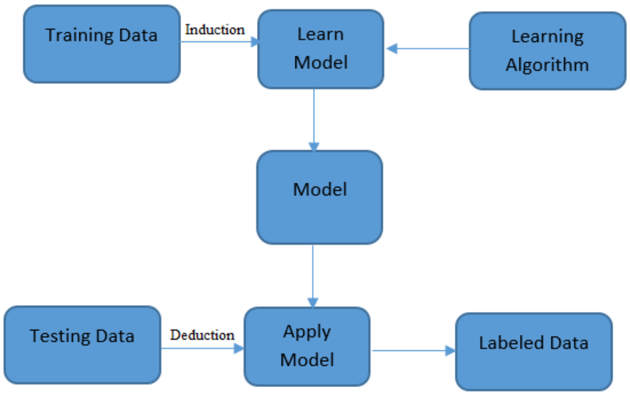

Fig. 3. The learning process: training and testing

Mathematical transformations are applied to signals to obtain further information from that signal that is not readily distinguishable in the raw signal. Most of the signals in practice are time-domain signals in their raw format. That is, whatever that signal is measuring, is a function of time. For example, looking at an EEG signal (Electroencephalogram, graphical recording of brain electrical activity), the typical shape of a healthy EEG signal is well known to neurophysiologist. Any significant deviation from that shape is usually considered to be a symptom of a pathological condition. This pathological condition, however, may not always be quite obvious in the original time-domain signal. Neurophysiologists usually use the time-domain EEG signals to analyze EEG. This, of course, is only one simple example why frequency content might be useful. The developed methodology is shown in Figure 4.

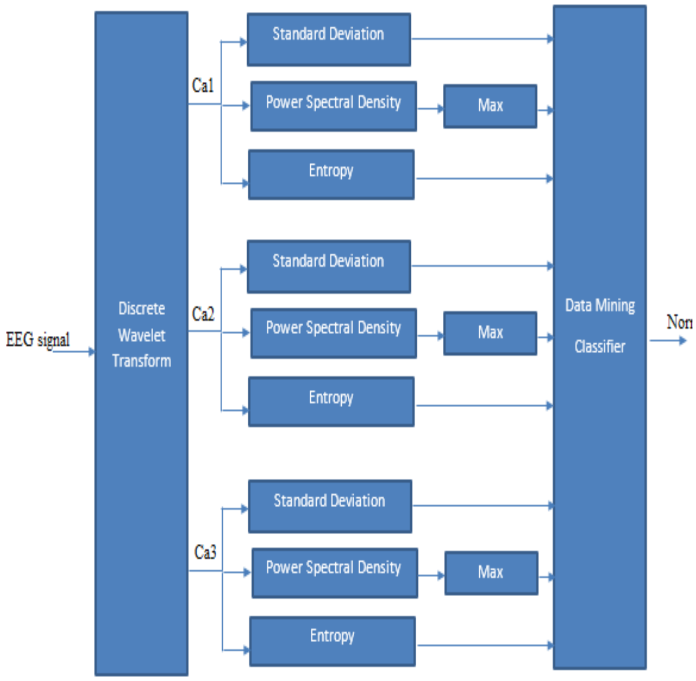

Fig. 4. The developed methodology

Wavelet analysis represents the next logical step: a windowing technique with variable-sized regions. Wavelet analysis allows the use of long time intervals where we want more precise low-frequency information, and shorter regions where we want high-frequency information. To make wavelet transformation, the time-domain signal from various high pass and low pass filters, which filters out either high frequency or low frequency portions of the signal. This procedure is repeated, every time some portion of the signal corresponding to some frequencies being removed from the signal [9-14]. The DWT is considerably easier to implement when compared to the CWT. The basic concepts of the DWT will be introduced in this section along with its properties and the algorithms used to compute it. In the discrete case, filters of different cutoff frequencies are used to analyze the signal at different scales. The signal is passed through a series of high pass filters to analyze the high frequencies, and it is passed through a series of low pass filters to analyze the low frequencies.

Calculating wavelet coefficients at every possible scale is a fair amount of work, and it generates a lot of data. Then the analysis will be much more efficient and just as accurate, then discrete wavelet transform (DWT) is obtained [14-18].

Given a signal S of length N, the DWT consists of log2N stages at most. Starting from S, the first step produces two sets of coefficients: approximation coefficients ca1, and detail coefficients cd1. These vectors are obtained by convolving S with the low-pass filter Lo_D for approximation, and with the high-pass filter Hi_D for detail, followed by dyadic decimation. The length of each filter is equal to 2N. If n = length (S), the signals F and G are of length n + 2N - 1, and then the coefficients ca1 and cd1 are of length is (n-1)/2 + N

The next step splits the approximation coefficients ca1 in two parts using the same scheme, replacing S by ca1 and producing ca2 and cd2, and so on.

In general case the decomposition step for one-dimensional DWT is in the block diagram shown below. It can be seen that for level J+1 the input signal S = ca J, then after low pass and high pass filtering the signal is down sampling by 2 to produce ca J+1 at the output, So the wavelet decomposition of the signal S analyzed at level j has the following structure: [caj, cdj, ..., cd1].

The standard deviations are calculated for each coefficient selected above. Where the standard deviation is

$$STD = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i-\mu)^2}$$

Power spectral density (PSD) refers to the amount of power per unit (density) of frequency as a function of the frequency. The power spectral density of each wavelet coefficient ca1, ca2, and ca3 is calculated, then the maximum value of the resulted power spectral density is achieved in order to be used as feature.

Another factor is used in this paper is entropy. The entropy of signal is a statistical measure of the randomness of the signal. The entropy of each beat was first calculated with two commonly used types of the entropy, Shannon entropy and log-energy entropy. The entropy is defined as followed

$$H = -\sum(\rho(x)\log(\rho(x))$$

The used methods are the Independent Component Analysis, Artificial Neural Networks, Support Vector Machines, Decision Trees, and Naïve Bayes, which will be used later in the experimental part.

Independent Component Analysis transform the data x into maximally independent variables. This method was proposed by Jeanny Hérault and

Bernard Ans. Basically, this method uses the equation s = Wx, whereas W is linear static transformation, x is the original data, and s is the hidden components.

Artifical Neural Network adopts the concept of biological neural networks. The artificial neural network is composed of a set of connected artificial neurons. Each connection has a weight, and the output of each neuron is a nonlinear function. The network is composed of three main layers: input, hidden, and output layers. The weights adjust according to the training data prediction error. The final weights are used in the classification of testing records. Recently, the number of hidden layers was increased, which introduced the deep learning topic. This improvement in the structure of the network contributed in increasing the accuracy of the classifier. Deep learning requires large amount of data to provide the optimal performance.

Naïve Base Classifier is a classification algorithm that relies on conditional probability. It utilizes Bayes' theorem and assumes a strong independence between features. It computes the occurrences of the features and the relation between them in the corpus considered. NB is good in practice even when the dimensionality of the input is high. NB is a good classifier for TC in general.

Decision Trees uses the training data to build a tree model that can be used later for classification purposes. Currently, many decision tree algorithms exist including Random Forest, Random Tree, J48, and CART [19,20].

3 Experiments and Results

This chapter describes the experimental work that was performed to distinguish between normal and abnormal EEG signals. The first section of this chapter introduces the data utilized in this paper and it is characteristics. The second section describes the first technique, which is the Wavelet transform of the EEG signals. The third section describes the second technique used in this work, which consists of Wavelet Transform and the phase space for the wavelet coefficients. The fourth section provides another analysis technique that depends only on the Phase Space plots without any transformation. The fifth section employs another analysis technique which is based on Fourier analysis only. The final section of this chapter provides a comparison among the different used techniques and their counter parts in the literature and discussion for all considered techniques.

3.1 Data

The datasets used in this paper are selected from the Epilepsy center in Bonn, Germany by Ralph Andrzejak. Five data sets containing quasi-stationary, artifact, e.g., due to muscle activity or eye movements, free EEG signals both in normal subjects and epileptic patients. Two sets denoted A, E are used in this work. Each set contains 100 single channel EEG segments of 23.6-sec duration. These segments were selected and cut out from continuous multichannel EEG recordings. Set A consisted of segments taken from surface EEG recordings that were obtained from five healthy volunteers using a standardized electrode placement. Volunteers were relaxed in an awake state with eyes open. Set E only contained seizure activity. All EEG signals were recorded with the same 128-channel amplifier system, using an average common reference (omitting electrodes containing pathological activity) E or strong eye movement artefacts. After 12 bit analogue-to-digital conversion, the data were written continuously onto the disk of a data acquisition computer system at a sampling rate of fs = 173.61 Hz. In all techniques 20 normal and abnormal single channel EEG segments are selected.

3.2 Experiments

Discrete wavelet transform DWT is used to decompose the EEG signal of each band of coefficients of normal and abnormal signals. DWT is taken of EEG signal for 3 levels using daubchies wavelet based function db4 that is the most common because of it is favourable characteristics, such as orthogonal and filters length that can be determined as the work needed. To determine the coefficients of each signal, wavelet coefficients ca1, ca2, and ca3 with frequency ranges of 0-fs//2, 0-fs/4, 0-fs/8, respectively are selected. These ranges contain the frequencies of EEG rhythms alpha, beta, theta, and delta. After that the statistical classification and power spectral density are taken to make classification of EEG signals.

The values of the standard deviation of each coefficient of normal and abnormal signals are calculated then power spectral density will be calculated.

Figures 5-12 display the power spectral density of each EEG signal and its wavelet coefficient ca1, ca2, and ca3. The figures of the abnormal case show that the spikes appear in the range of 1-4 Hz with maximum power (maximum amplitudes), but they don't appear in normal [20].

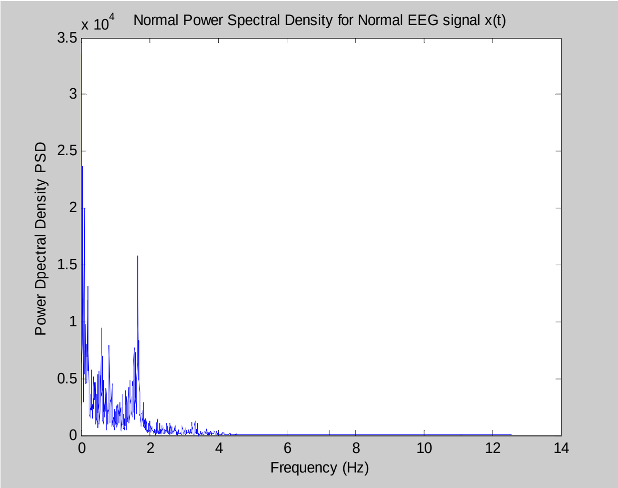

Fig.5 PSD of normal EEG signal

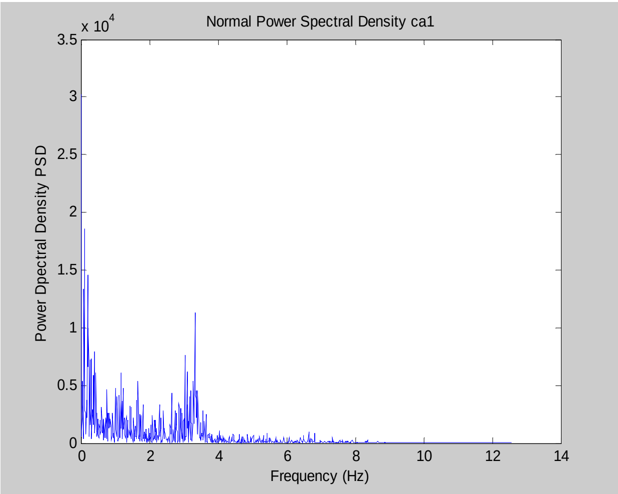

Fig.6 PSD of normal ca1

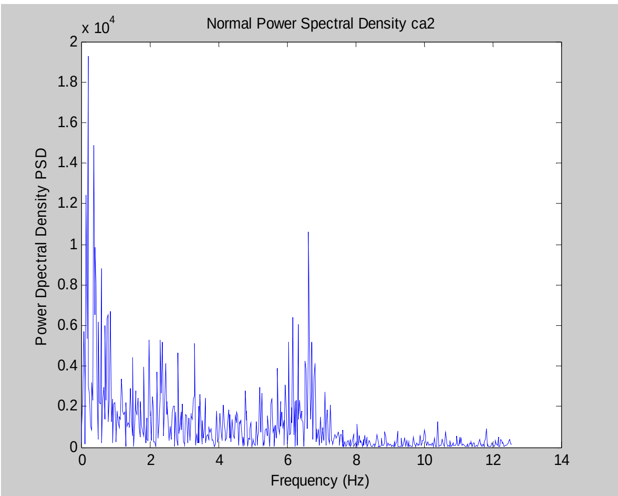

Fig.7 PSD of normal ca2

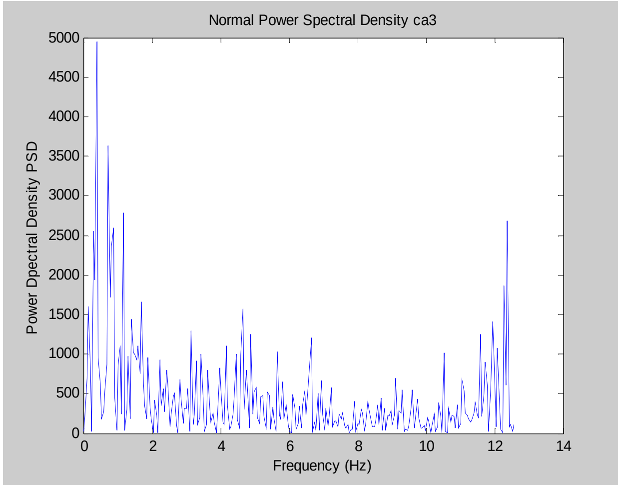

Fig.8 PSD of normal ca3

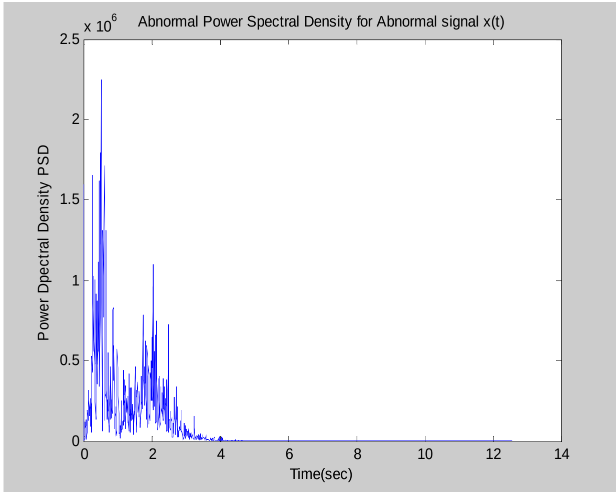

Fig.9 PSD of abnormal signal EEG

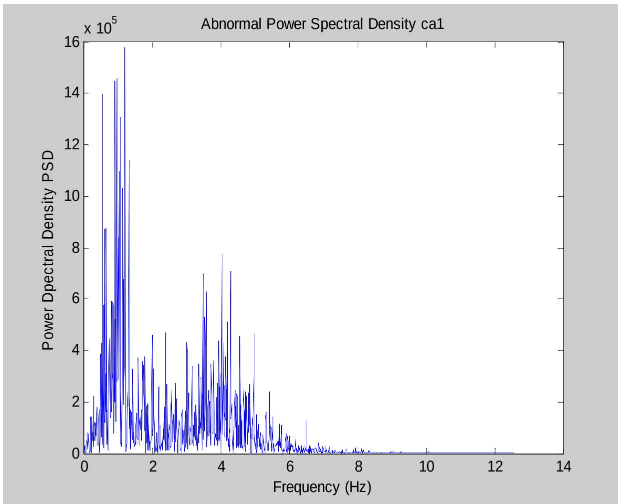

Fig.10 PSD of abnormal ca1

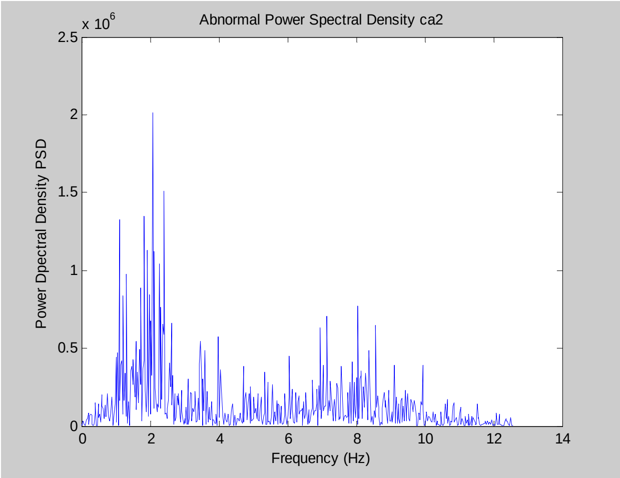

Fig.11 PSD of abnormal ca2

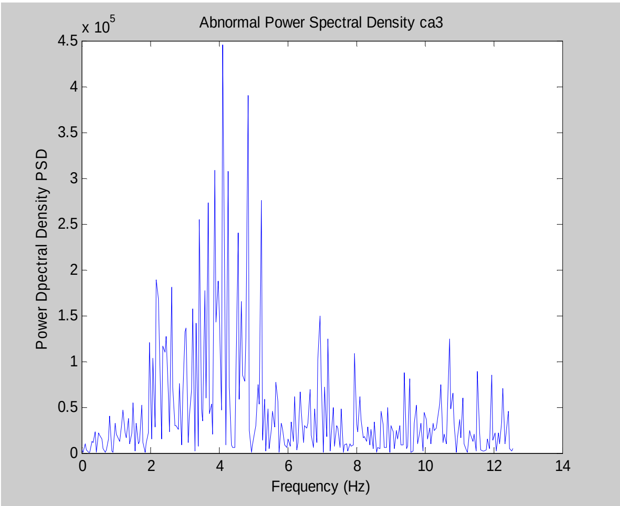

Fig.12 PSD of abnormal ca3

The total number of signals used are 150 signals. As for the classifiers, we used five classifiers from the literature. Namely, we compared SVM, ANN, Naïve Bayes, and Decision Trees. As for the implementation of these classifiers, we used WEKA toolkit.

Weka is a popular suite of machine learning software written in Java, developed at the University of Waikato, New Zealand. Weka is a free software tool available under the GNU General Public License. It contains a collection of visualization tools and algorithms for data analysis and predictive modelling that support data pre-processing, clustering, classification, regression, visualization, and feature selection. Weka has a powerful graphical user interface that supports its functionality.

Regarding the evaluation measurements, we used training time, classification time, and F1 measurement, they are defined as follows.

- Training time: It is the time needed to train the classifier and build the model.
- Testing time: It is the time needed to classify the testing instances.
- F1 Measurement: It is an indication of the classification accuracy. It is defined as the harmonic mean of the recall and the precision. Its formula is given as follows.

$$F1 = \frac{2 * R * P}{R + P},$$

Whereas R presents the recall, which is the percentage of retrieved relevant and P presents the precision. The formulas of recall and precision are given as follows.

$$R = \frac{TP}{TP+FN},$$

$$P = \frac{TP}{TP+FP},$$

Whereas TP is true positives, FP is the false positive From Table 1, it is obvious that the use of ICA proved its high performance in EEG signal classification. This is reasonable as ICA is capable of eliminating noise from images and concentrating on important features only. Furthermore, Artificial neural network classifier proved to be of superior accuracy, but its training and testing time is slightly larger than other classifiers. On the other hand, Naïve Bayes proved to be a very fast classifier that works efficiently with big datasets, despite its average accuracy compared with other classifiers. Therefore, the choice of the classifier is domain dependent, as certain domains seek the optimal accuracy while others seek fast execution time.

Table 1 Performance Measurement

Classifier	Training Time	Classification Time	F1 Measurement
SVM	1.05	10	0.705
Artificial Neural Networks	26.6	240	0.909
Decision Trees (J48)	0.39	1	0.76
Naïve Bayes	0.03	1	0.837

10-fold cross validation has been used in this paper. In 10-fold the training set will be randomly splitted into 10's that have approximately the same size. Then the classifier will be trained using (8) subsets. One of the two remaining subsets will be used for validation and the last for testing.

This process will be repeated 10 times, while a different subset is used for testing and validation. Table 2 shows the average accuracy for 2-Fold, 4-Fold, 6-Fold, 8-Fold and the 10-Fold cross-validation for Naïve Bayes classifier as an example.

Table 2: Accuracy for 2-Fold, 4-Fold, 6-Fold, 8-Fold and the 10-Fold cross-validation

.K-Fold	Accuracy (%)
2-Fold	91.5044 %
4-Fold	93.2743 %
6-Fold	93.2743 %
8-Fold	92.3894 %
10-Fold	99.4819 %

As the number of features in that matrix was 9 features, we used PCA dimensionality reduction method and we selected around 77% of the features (7 features) with the highest importance.

There are many different reduction techniques available including Principal Component Analysis (PCA), and Chi-Squared for feature selection and reduction. We used the Describable EEG signal Dataset to compare PCA and Chi-Squared on a reduced number of dimensions varying between 1 and 9 dimensions. Fig. 13 depicts the F1 rating of these two methods. Clearly, PCA F1 measurement of PCA outperformed the other method.

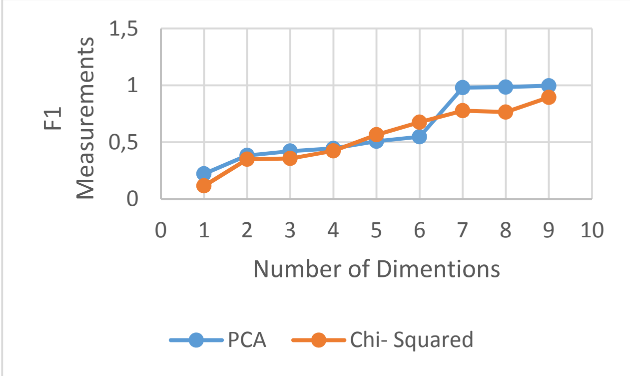

Fig.13 F1 Measurement for PCA and Chi-Squared

4 Conclusions

This paper presents a framework for analysing EEG time-series data and demonstrates its use for a practical problem: detecting seizures from the intracranial EEG (seizure analysis). The methodology of this paper used a modified technique in analysing EEG signal to distinguish between normal and abnormal (epileptic seizures). In this technique disorder in the brain is easily detected using simple classifier depends on data mining techniques, which uses the features, standard deviation STD, maximum of PSD for DWT coefficients ca1, ca2, and ca3, and the entropy. the technique detected the spikes caused by epilepsy and determined easily their locations (i.e. frequency ranges) using simple approach.

References

[1] www.bEEG-MEG
[2] Szilagyi L., Z.benyo, S.M. Szilagyi "A new method for epileptic waveform recognition using wavelet decomposition and artificial neural Networks", IEEE trans. Houston, TX, USA, October 2002 .
[3] Barreto A., N. Chin., J. Andrian J. Riley " Multiresolution Characterization of Interictal Epileptic Spikes based on a Wavelet Transformation". University Park ,Miami, Florida, 33199
[4] Gotman .J, 1985 "Seizure recognition and analysis," Electroencephalogram Clinical Neurophysiology Suppl,vol.37, pp. 133-45,.
[5] Zijlmans. M, Flanagan .D, and Gotman .J, 2002,"Heart rate changes and ECoG abnormalities during epileptic seizures: prevalence and definition an objective clinical sign," Epilepsia, vol. 43, pp. 847-54,

[6] Theodore .H and Haas .L, 1991, "A personal computer controlled long-term video-EEG monitoring, seizure detection, and collection system,"American Journal of EEG Technology, vol. 31, pp. 103-108,.

[7] Zaveri .H, Williams .W, and Sackellares .J, 1993, "Energy based detection of seizures", presented at 15th Annual International Conference on Engineering and Medicine in Biology.

[8] Schiff .J, Aldroubi .A, Unser .M, and Sato .S, 1994, " Fast Wavelet Transformation of EEG Electroencephalography and Clinical Neurophysiology", 91, 442 .

[9] Williams .W, Zaveri .H, and Sackellares .J, 1995 "Time-Frequency Analysis of Electrophysiology Signals in Epilepsy," IEEE Engineering in Medicine and Biology, pp. 133-143,.

[10] Benedetto .J and Colella .D, 1995. "Wavelet analysis of spectrogram seizure chirps,"presented at Wavelet Applications in Signal and Image Processing III",

[11] Barreto .A, Chin .N, Andrian .J. Riley .J, " Multiresolution Characterization of Interictal Epileptic Spikes based on a Wavelet Transformation". University Park, Miami, Florida, 33199.

[12] szilagyi .L, benyo .Z, Szilagyi .S, 2002 "A new method for epileptic waveform recognition using wavelet decomposition and artificial neural networks", IEEE trans. Houston, TX, USA,

[13] Chiri Yamaguchi, 2003, "Fourier and Wavelet Analyses of Normal and Epileptic
Electroencephalogram (EEG)" Proceedings of the 1st International IEEE EMBS Conference on Neural Engineering Capri Island, Italy.

[14] Arnaud Delorme and Scott Makeig, 2003 "EEG Changes Accompanying
Learned Regulation of 12-Hz EEG Activity", IEEE, VOL. 11, NO. 2,

[15] Quiroga .R, Nadasdy .Z Ben-Shaul .Y, 2004 LETTER Communicated by Maneesh Sahani "Unsupervised Spike Detection and Sorting with Wavelets Superparamagnetic Clustering Neural Computation" 16, 1661– 1687 Massachusetts Institute of Technology.

[16] Gardner .A, 2004"A Novelty Detection Approach to Seizure Analysis from Intracranial EEG". A Dissertation Presented in Doctor of Philosophy to the academic faculty in Electrical Engineering School of Electrical and Computer Engineering “Georgia Institute of Technology” Atlanta, GA USA

[17] Gotman .J, "Automatic seizure detection: improvements and evaluation,"
Electroencephalogr Clin Neurophysiol, vol. 76, pp. 317-24, 1990.

[18] Qu .H and Gotman .J, 1993, "Improvement in seizure detection performance by automatic adaptation to the EEG of each patient," Electroencephalogr Clin Neurophysiol, vol. 86, pp. 79-87,

[19] Qu .H and Gotman .J, 1995, "A seizure warning system for long-term
epilepsy monitoring," Neurology, vol. 45, pp. 2250-4,.

[20] Qu .H and Gotman .J, 1997,"A patient-specific algorithm for the detection of seizure onset in long-term EEG monitoring: possible use as a warning device," IEEE Trans Biomed Eng, vol. 44, pp. 115-22,.