EEG Based Epilepsy Diagnosis System using Reconstruction Phase Space and Naïve Bayes Classifier

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

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

Abstract: - Electroencephalogram (EEG) is one the most used tools for the diagnoses and analysis of epilepsy. The diagnosis of epilepsy diseases are still made by physicians manually. This process is time consuming and subjective. In this study, EEG signal is analyzed by Discrete Time Wavelet Transform and Reconstruction Phase Space. Both techniques are used together to extract EEG features that allows Naïve Bayes classifier to diagnose the epilepsy diseases and classify the corresponding EEG signals into "normal" or "abnormal" classes based on the extracted features. To assess the performance of the proposed system, we conducted a simulation experiment that involved 200 EEG signals from publicly available EEG dataset from University of Bonn. The proposed algorithm shows excellent accuracy compared with other techniques.

Key-Words: - Epilepsy, Naïve Bayes, Genetic Algorithm, Phase Space, Discrete Time Wavelet Transform.

1 Introduction

The electrical activity of active nerve cells in the brain produces currents spreading through the head. These currents reach the scalp surface, and the resulting voltage differences on the scalp can be recorded as the electroencephalogram (EEG). EEG reflect brain electrical activity.

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. 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. Well-known causes of epilepsy may include: genetic disorders, traumatic brain injury, metabolic disturbances, alcohol or drug abuse, brain tumor, stroke, infection, and cortical malformations.

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.
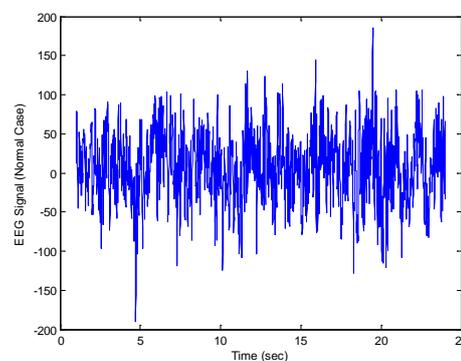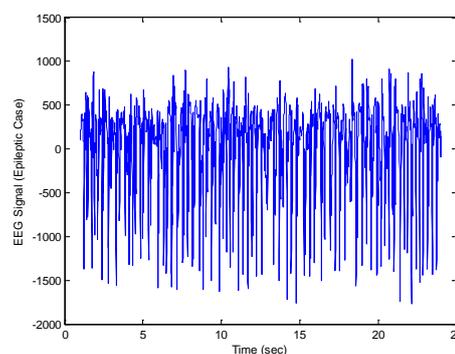

Fig. 1.Normal EEG Signal

Fig. 2. Epileptic EEG Signal

In this paper EEG signal is analyzed by optimized phase space and naïve byes classifier to detect the epilepsy. The proposed approach offers excellent detection technique.

EEG signal classifiers play a particularly important role in EEG signal processing. Many recent studies have proposed new methods for classifying and detecting Epilepsy [1-12]. Still more research efforts are seriously needed for clinical implementation. The proposed methodology in this paper investigates a new epilepsy diagnosis method based on Discrete Wavelet Transform (DWT), Reconstruction Phase Space (RPS) and Naïve Bayes classification.

The paper is organized as follows: in section 2 the methodology of the work is described. In Section 3 the proposed methodology is applied to real EEG signals some of them suffer from epileptic seizures. Finally, the paper is concluded in section 4.

2. METHODOLOGY

Mathematical transformations are applied to signals to obtain hidden information from that signal that is not readily distinguishable in the raw signal. Looking at an EEG signal, 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. This makes frequency content useful. The basic architecture of the proposed method is shown in Fig 3.
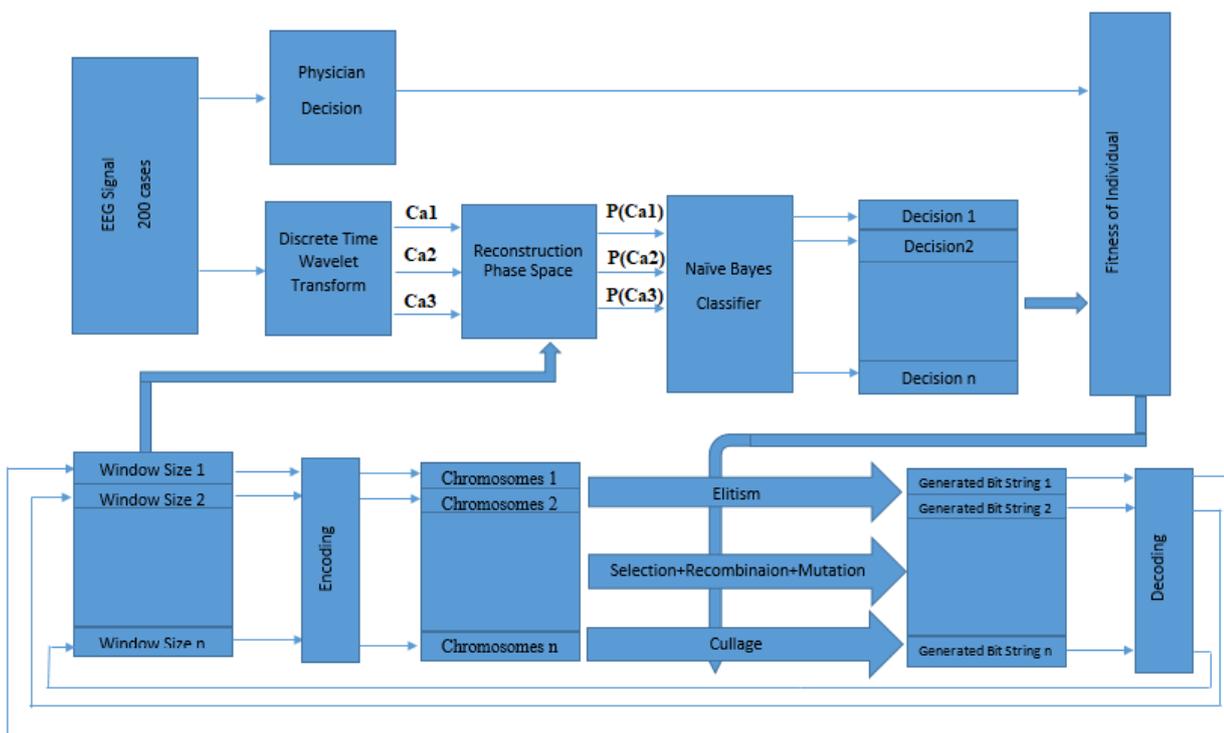

Fig. 3. The Developed System Methodology

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 favorable 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 phase space is then reconstructed for all series wavelet coefficients.

The abnormal ca's show wide range of scattering and deviation when compared with the phase space plots of normal ca's. This observation is that it enables us to come up with a simple feature extraction method to distinguish normal from abnormal signals. The resulted features will be classified by Naïve Bayes classifier as normal or abnormal cases. Genetics algorithm will be used to optimize Phase Space unit in order to select the best window size and type used in feature selection.

2.1. Discrete Time Wavelet Transform (DTWT) Unit

Wavelet analysis represents a windowing technique with variable-sized regions. The wavelet transform can provide a time frequency description of signal. To make wavelet transformation, the time-domain signal will be filtered out either by removing high frequency or low frequency portions of the signal. This procedure is repeated, every time some portion of the signal corresponding to some frequencies are removed from the signal.

Given a signal EEG signal of length N, the DWT consists of $\log_2 N$ stages at most. Starting from EEG signal, the first step produces two sets of coefficients: approximation coefficients ca1, and detail coefficients cd1. These vectors are obtained by convolving EEG signal with the low-pass and the high-pass filters as shown in Fig. 4. 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 [12].The next step splits the approximation coefficients ca1 in two parts using the same scheme replacing EEG signal by ca1 and producing Ca2 and Cd2, and so on. It can be seen that for level i+1 the input signal after low pass and high pass filtering the signal is down sampling by 2 to produce Cai+1 at the output. For example the wavelet decomposition of the signal EEG signal analyzed at level i has Cai+1, cdi+1
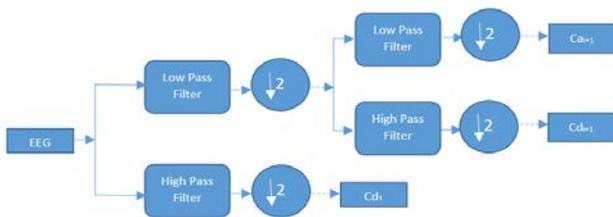

Fig. 4. Diagrams of DTWT EEG Signal Decomposition

2.2. Reconstructed Phase Space (RPS) Unit

The reconstructed phase space (RPS) technique has been used to study the nonlinear dynamical behavior of a variety of time series analysis. The reconstructed phase space is an example of higher dimension transformation of the time series, where the RPS is an n dimensional space in which a signal is plotted against time-delay versions of itself. Each point in such a phase space is calculated as follow:

$$X_n = \begin{bmatrix} x_{n-(d-1)\tau} & ..... & x_{n-\tau} & x_n \end{bmatrix},$$
$$n = (1+(d-1)\tau).....N \qquad (1)$$

Where N is the dimension of the time series, τ is the time delay, and d is the embedding dimension. Then the entire phase space is generated by:

$$X = \begin{bmatrix} X_{1+(d-1)\tau} \\ X_{2+(d-1)\tau} \\ \vdots \\ X_N \end{bmatrix} = \begin{bmatrix} x_{1+(d-1)\tau} & \cdots & x_{1+\tau} & x_n \\ x_{2+(d-1)\tau} & \cdots & x_{2+\tau} & x_2 \\ \vdots & & \ddots & \\ x_N & \cdots & x_{N-(d-2)\tau} & x_{N-(d-1)\tau} \end{bmatrix} \quad (2)$$

The trajectory matrix is formed by compiling its row vectors from the vectors that are created by equation 1. The resulted matrix is a mathematical representation of the reconstructed phase space.

To reconstruct the phase space the time lag and the dimension should be determined, the determination of these parameter affects the classification method, so they can be determined to obtain the maximum accuracy of the classification. The time lag can be determined by using the first minimum of the auto-mutual information function method, the first zero crossing of the autocorrelation, or empirically such as to obtain maximum classification accuracy. The dimension can be selected using the false nearest neighbors, Cao's method, or by an empirical method to obtain maximum accuracy.

The choice of the lag time will affect the distribution of the attractor in the reconstructed phase space. For a nonlinear periodic time series, a first minimum of the auto-mutual information function can be used to determine the optimal time delay for embedding the signal in the phase space reconstruction. Mutual information function provides a measure for dependency within and between time series. In other words, when taking a sequence of time series measurements over a

specified period of time, the uncertainty in the prediction of the next measurement can be estimated.

Let us consider two time series X and Y, the uncertainty $I_{XY}(x_i, y_j)$ which is called the mutual information about measurements x_i given y_j is expressed by:

$$I_{XY}(x_i, y_j) = \log_2\left(\frac{P_{XY}(x_i, y_j)}{P_X(x_i)P_Y(y_j)}\right) \quad (3)$$

Where $P_X(x_i)$ is the probability of observing x_i in X time series, $P_Y(y_j)$ is the probability of observing y_j in Y time series and $P_{XY}(x_i, y_j)$ is the joint probability of observing x_i and y_j in the two time series. The average mutual information of the two time series is the mean of mutual information over all of the measurements in the two series, which is given by:

$$I_{XY}(T) = \frac{1}{N}\sum_{ij} P_{XY}(x_i, y_j)I_{XY}(x_i, y_j) \quad (4)$$

Now, consider one time series X with period of T, and then the average mutual information of that series can be calculated as discussed above. Mathematically, the average auto-mutual about a measurement $x_{t+\tau}$ given a measurement x_t at time t over all the time t is given by:

$$I(\tau) = \frac{1}{N}\sum_{t=1}^{T} P(x_t, x_{t+\tau})\log_2\left[\frac{P(x_t, x_{t+\tau})}{P(x_t)P(x_{t+\tau})}\right] \quad (5)$$

The mutual information function applied on a time series is similar to the autocorrelation function in which the degree of dependency is measured. Mutual information function measures the dependency even the time series is nonlinear or chaotic. The advantages of autocorrelation are that it can be calculated quickly and it does a good job of describing a linear system. The disadvantage is that the autocorrelation function always assumes that the underlying process is linear and calculates a value for dependency; if the underlying process is not linear, the value would be incorrect.

In the reconstructed phase space, each specific coefficient resulted from Discrete Time Wavelet Transform unit ca1, ca2 and ca3 has its geometrical distribution, so identification parameters can be chosen depending on the distribution of data within RPS. A simple feature extraction method has been used in this paper to extract the features from ca1, ca2 and ca3. The proposed method uses a box a centered at zero and spans a range between -100 to 100. This box measures the clustering of points around the center of the phase space of ca1, ca2 and ca3. The percentages of the number of points bounded by box (window) is calculated for all wavelet coefficients as:

$$P(.) = \frac{N(.)}{N_T(.)} * 100\% \quad (6)$$

Where N (.)=Number of points in the selected area and N_T(.)=Total number of points in the RPS. P(.) will be used as feature vector to distinguish normal from abnormal signals in Naïve Bayes Classifier.

2.3. Genetic Algorithm (GA) Unit

Genetic Algorithm (GA) is one of the most popular derivative free optimization techniques which is based on the principles of evolution and natural genetics [13].

Genetic algorithms make multiple way search by creating a population of candidate solutions instead of just test one single solution. A starts by constructing a new population using genetic operation such as crossover and mutation through an iterative process until some convergence criteria are met.

The resulted new population will be decoded back to its original format. A new generation is created by repeating the selection, recombination and mutation processes until all chromosomes in the new population replace the initial population.

The main step after creating a population of chromosomes is to calculate the fitness value of each member in the population. The fitness value fi of the ith weight parameter is the objective function evaluated at this weight set. The fitness function is chosen to be the root mean squared differences between the correct decision specified by physician T and the decision given by the Naïve Bayes Classifier $\hat{T}$. The Root Mean Squared Error (RMSE) is given by equation (7). Root Mean Square Error (RMSE) is the standard deviation of the residuals (prediction errors). Residuals are a

measure of how far from the actual class labeled by physician,

$$\textbf{\textit{Objective Function}} = \textbf{\textit{RMSE}} = \sqrt{\frac{\sum_{i=1}^{n}(T-\hat{T})^2}{n}} \quad (7)$$

Where n is number of records of training data

By this definition, then, the lower the fitness, the better the developed model, and a fitness of zero means that the model achieves the desired behavior for all inputs.

When we are sampling we should try different window. The size of the window should be a compromise between two requirements:

- It is large enough to contain sensible information even though the signal is sampled and to retain the spectral properties of the original signal.
- It is small enough to introduce generality in the data. To find any general properties in signal(s), we have to restrain to local neighborhoods.

The most usual windowing method is the rectangular one, where a sample window of size m x m with top corner positioned at x0, y0 is obtained by multiplying the original phase space output function i(x,y) with window function

$$w(x,y,x_0,y_0) = \begin{cases} 1 \; if \; x_0 \le x \le x_0 + m - 1 \; and \; y_0 \le y \le y_0 + m - 1 \\ 0 \; otherwise \end{cases}$$

$$(8)$$

In order to select a window dimension, different box sizes and types were chosen, then the accuracy of each box was calculated to get the optimal one using GA.

2.4. Naïve Bayes Classifier Unit

Naïve Bayes classifiers still tend to perform very well under this unrealistic assumption. Naïve Bayes is one of the most effective and efficient classification algorithms [14].

The probability that a record with feather vector F={$F_1, F_2, ..., F_n$} belongs to class Ci, Where $C_i \in \{C_1, C_1, ..., C_k\}$

$$p(C_i|\textbf{\textit{F}}) = p(C_i|F_1, F_2, ..., F_n) = \frac{p(F_1, F_2, ..., F_n|C_i).P(C_i)}{p(F_1, F_2, ..., F_n)}$$

$$(9)$$

- $p(C_i|\textbf{\textit{F}})$ is the posterior probability of class (C_i) given features ($F_1, F_2, ..., F_n$).
- $P(C_i)$ is the prior probability of class (C_i).
- $p(F_1, F_2, ..., F_n|C_i)$ is the likelihood which is the probability of features ($F_1, F_2, ..., F_n$) given class (C_i)..
- P(x) is the prior probability of features ($F_1, F_2, ..., F_n$).

Calculating $p(C_i|\textbf{\textit{F}})$ is the main aim in Naïve Bayes Classifier. Specifically, we want to find the value of C_i that maximize $p(C_i|F_1, F_2, ..., F_n)$. Since $p(F_1, F_2, ..., F_n)$ is constant number for all k values, then equation (10) can simply say

$$p(C_i|\textbf{\textit{F}}) = p(C_i|F_1, F_2, ..., F_n) \propto$$
$$(p(F_1, F_2, ..., F_n|C_i).P(C_i)) \quad (8)$$
Where

$$P(C_i) = \frac{N_{C_i}}{N_T} \quad\quad\quad (10)$$

- Nci is count of samples from class Ci.
- NT is Count of all samples.

Naive Bayes classifier assumes Conditional independence that the effect of the value of a feature (F) on a given class (C) is independent of the values of other predictors. Then the $p(F_1, F_2, ..., F_n|C_i)$ can be written as

$$p(F_1, F_2, ..., F_n|C_i) = \prod_{j=1}^{n} p(F_j|C_i) \quad (11)$$

Finally the Naive Bayes calculate the posterior probability for each class. Choose value of Ci that maximizes P(Ci | F1, F2, …, Fn) is equivalent to choosing value of Ci that maximizes P(F1, F2, …, Fn|Ci) P(Ci)

The class with the highest posterior probability is the predicted one. The estimated class $\hat{C}_i$ Corresponding to F is

$$The \; Predicted \; Class \; \hat{C}_i =$$
$$\max_i p(C_i|F_1, F_2, ..., F_n), 1 \le i \le K \quad (12)$$

3. EXPEREMENT AND RESULT

3.1. Experiment

The datasets used in this research are selected from the epilepsy center in Bonn, Germany by Ralph Andrzejak [15]. 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. In this current research, we only use two sets: A and E. Set A as "normal class" and set E as "epileptic class". 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). After 12 bit analog-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 this paper, 200 normal and abnormal single channel EEG segments are selected.

The discrete wavelet coefficients ca1, ca2, and ca3 will be calculated. Then the phase space of each band of coefficients is plotted for normal and abnormal signals. In order to reconstruct the phase space for time series signal, the lag time and the embedding dimension should be determined. The embedding dimension used in this technique is chosen empirically as 2. There are many ways to determine the time lag of the reconstructed phase space, one of these methods is the auto-mutual information approach. The optimal time delay is calculated using the first minimum of the auto-mutual information for each time series signal. The phase space is reconstructed for all series wavelet coefficients using a time lag of 2 and an embedding dimension of 2. Fig. 5 and Fig.6 shown below are examples of normal and abnormal phase space for ca1, ca2, and ca3. As seen from the figures below in the normal signal the majority of points in ca3 are centered on zero without scattering with small amplitudes compared to the points of the same coefficients in abnormal signal. This distribution is based on the fact that the abnormal signal has been affected by epileptic seizures that caused spikes with high amplitudes and low frequencies.
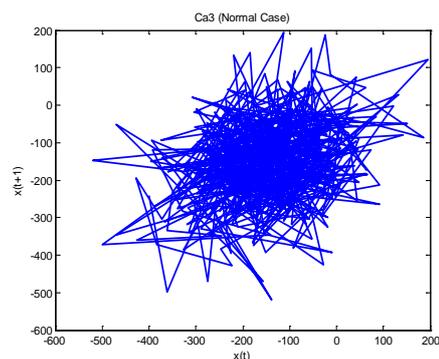

Fig.5. Phase Space Plot of Normal ca3
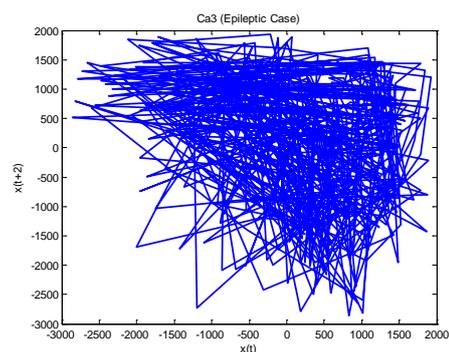
Fig.6. Phase Space Plot of Abnormal ca3

The results of only 40 cases out of 200 cases are shown in table 1. The experiments were performed with input window size of (-100×100) because through this size we have achieved the best classification accuracy using Genetic algorithm. Table 2 show the relation between rectangular and circular window size and the overall accuracy achieved by Genetic Algorithm. Fig.7 shows that the rectangular window of size 100x100 has the best accuracy compared with other windows types and sizes.

Table 1 Feature Extraction of 20 Cases (Normal and Abnormal).

	Normal Cases				Abnormal Epileptic Cases		
Case Num	P (ca1)	P (ca2)	P (ca3)	Case Num	P (ca1)	P (ca2)	P (ca3)
1	219	105	41	21	11	3	2
2	385	247	160	22	18	5	5
3	371	240	187	23	3	0	1
4	414	282	188	24	9	7	14
5	402	296	212	25	12	3	2
6	121	10	0	26	118	46	18
7	407	252	196	27	46	37	22
8	265	127	35	28	41	13	10
9	421	273	206	29	6	1	0
10	106	14	1	30	0	1	1
11	493	421	365	31	124	61	40
12	433	306	227	32	12	8	0
13	465	350	262	33	81	38	23
14	444	308	235	34	6	7	6
15	296	167	82	35	12	3	1
16	430	330	240	36	12	12	6
17	262	111	31	37	16	6	2
18	416	266	184	38	3	0	2
19	423	337	302	39	257	137	69
20	438	344	310	40	49	18	11

Table 2 Correct Classification Rate for Different Rectangular Windowing Size

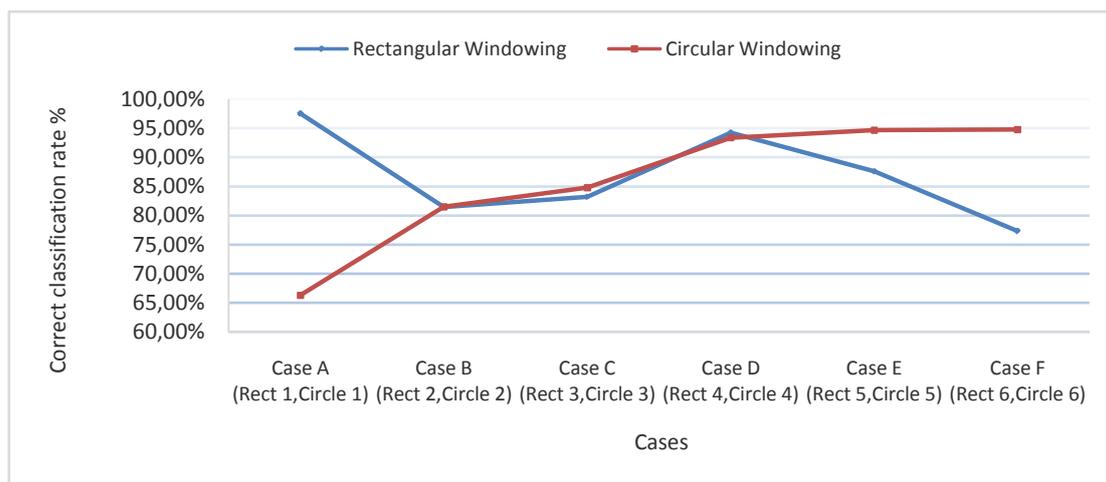| Window size	Correct classification rate (%)	Circular Window size (radius)	Correct classification rate (%)
-100 to 100	97.54%	25	66.3%
-50 to 50	81.46%	50	81.5%
-75 to 75	83.23%	75	84.8%
50 to 100	94.26%	100	93.4%
-50 to 0	87.63%	125	94.7%
-100 to -50	77.34%	150	94.8%

Fig.7. Accuracy of Different Rectangular and Circular Window Sizes

In this study, WEKA (Waikato Environment for Knowledge Analysis) [16] is used to construct Naïve Bayes Classifier according from training set. 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.

Once the features have been extracted and grouped into a feature vector, classification takes place, where each EEG signals are classified in one of the two classes: Normal or Abnormal.

3.2. Analysis of the Results

Fig. 8 Shows the Weka run information. To further validate the results, 10-fold cross validation was used. 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.

```
=== Run information ===
Scheme:    weka.classifiers.bayes.NaiveBaye
Instances:  200
Attributes: 5
Test mode:  10-fold cross-validation
=== Classifier model (full training set) ===
Naive Bayes Classifier
             Class
Attribute    normal abnormal
            (0.5)   (0.5)
================================
ID
  mean         95.5     105.5
  std. dev.  57.5174   57.5174
  weight sum   100       100
  precision     1         1
P (ca1)
  mean       72.3392   3.4683
  std. dev.  12.7675   4.4593
  weight sum   100       100
  precision   4.9547   4.9547
P (ca2)
  mean       55.384   0.8791
  std. dev.  15.5033   2.6373
  weight sum   100       100
  precision   4.3956   4.3956
P (ca3)
  mean       38.7057  0.4501
  std. dev.  15.6167   1.3502
  weight sum   100       100
  precision   4.5007   4.5007
```

Fig.8. Naïve Bayed Run Information

Fig. 9 shows the resulted Classification model evaluation results. The performance measurements used for this paper were precision, recall, F1 measure and accuracy. The proposed model shows

high accuracy on epilepsy classification (up to 100%) on the test data.

We examined agreement between the results generated by the developed system and the one by the physicians. We constructed the confusion matrix for each class (present or not -present). The confusion matrix has the form shown in Fig.9,

```
Time taken to build model: 0 seconds

=== Stratified cross-validation ===
=== Summary ===

Correctly Classified Instances      200         100    %
Incorrectly Classified Instances    0           0      %
Kappa statistic                 1
Mean absolute error             0
Root mean squared error         0
Relative absolute error         0.0004 %
Root relative squared error     0.0021 %
Total Number of Instances       200

=== Detailed Accuracy By Class ===

        TP Rate  FP Rate Precision Recall F-Measure MC(
        1.000    0.000   1.000     1.000  1.000     1.000
        1.000    0.000   1.000     1.000  1.000     1.000
Weighted Avg.  1.000  0.000  1.000  1.000  1.000    1.000

=== Confusion Matrix ===

  a    b  <-- classified as
100   0 |  a = normal
  0 100 |  b = abnormal
```

Fig.9. Epilepsy Classification Results

Table 3 presents summary of EEG-based epilepsy diagnosis that used the same dataset (set A vs set E). The summary has been adapted by Al harabi et al in [11]. Our work achieved perfect classification performance compared with other methods as shown in Table 3.

4. CONCLUSION

In this paper we have extracted EEG features from the input EEG signals using both Discrete Time Wavelet Transform and Reconstruction Phase Space. We have shown that the develop methodology that uses Naïve Bayes can classify EEG signal as normal or abnormal (epileptic seizures) in the same level of accuracy as that of human evaluator.

Table 3 Different EEG- Based Epilepsy Diagnosis

Author	Feature Extraction	Classifier	Accuracy (%)
Niqam et al [17]	Non-linear filter	ANN	97.2
Kannathal et al [18]	Entropies	ANFIS	92.2
Subasi et al [19]	DWT	Mixture of experts	94.5
Srinivasan et al [20]	Approximate entropy (Apen)	Elman ANN	100
Ocak [21]	A pen on DWT	ANN	96
Dhiman et al [22]	DWT, GA-SVM	SVM	
Khalil Al Sharabi et al [11]	DWT, Shannon entropy	ANN	100
This work	DWT, Reconstruction phase Space	Naïve Bayes	100

References:

[1] Quintero-Rincón, J. Prendes, M. Pereyra, H. Batatia and M. Risk, "Multivariate Bayesian classification of epilepsy EEG signals," 2016 IEEE 12th Image, Video, and Multidimensional Signal Processing Workshop (IVMSP), Bordeaux, 2016, pp. 1-5.

[2] E. B. Assi, D. K. Nguyen, S. Rihana and M. Sawan, "Refractory epilepsy: Localization, detection, and prediction," 2017 IEEE 12th International Conference on ASIC (ASICON), Guiyang, China, 2017, pp. 512-515.

[3] Y. Wang, Z. Li, L. Feng, H. Bai and C. Wang, "Hardware design of multiclass SVM classification for epilepsy and epileptic seizure detection," in IET Circuits, Devices & Systems, vol. 12, no. 1, pp. 108-115, 1 2018.

[4] Z. Lasefr, S. S. V. N. R. Ayyalasomayajula and K. Elleithy, "Epilepsy seizure detection using EEG signals," 2017 IEEE 8th Annual Ubiquitous Computing, Electronics and Mobile Communication Conference (UEMCON), New York City, NY, USA, 2017, pp. 162-167.

[5] Z. Lasefr, R. R. Reddy and K. Elleithy, "Smart phone application development for monitoring epilepsy seizure detection based on EEG signal classification," 2017 IEEE 8th Annual Ubiquitous Computing, Electronics and Mobile Communication Conference (UEMCON), New York City, NY, USA, 2017, pp. 83-87.

[6] H. Rajaguru and S. K. Prabhakar, "Optimization of fuzzy outputs for classification of epilepsy from EEG signals using linear discriminant analysis," 2017 International conference of Electronics,

Communication and Aerospace Technology (ICECA), COIMBATORE, India, 2017, pp. 598-602.

[7] H. Rajaguru and S. K. Prabhakar, "Modified expectation maximization based sparse representation classifier for classification of epilepsy from EEG signals," 2017 International conference of Electronics, Communication and Aerospace Technology (ICECA), COIMBATORE, India, 2017, pp. 607-610.

[8] H. Rajaguru and S. K. Prabhakar, "A comprehensive analysis of support vector machine and Gaussian mixture model for classification of epilepsy from EEG signals," 2017 International conference of Electronics, Communication and Aerospace Technology (ICECA), COIMBATORE, India, 2017, pp. 585-593.

[9] R. Dogaru and I. Dogaru, "A low complexity solution for epilepsy detection using an improved version of the reaction-diffusion transform," 2017 5th International Symposium on Electrical and Electronics Engineering (ISEEE), Galati, 2017, pp. 1-6.

[10] S. Supriya, S. Siuly and Y. Zhang, "Automatic epilepsy detection from EEG introducing a new edge weight method in the complex network," in Electronics Letters, vol. 52, no. 17, 2016, pp. 1430-1432.

[11] K. AlSharabi, S. Ibrahim, R. Djemal and A. Alsuwailem, "A DWT-entropy-ANN based architecture for epilepsy diagnosis using EEG signals," 2016 2nd International Conference on Advanced Technologies for Signal and Image Processing (ATSIP), Monastir, 2016, pp. 288-291.

[12] Mohammad A. Obaidat and Ayman M. Mansour," EEG signal Analysis using Decomposed Wavelet Transform and Fuzzy Inference System", International Journal of Electrical, Electronics and Computer Systems. Vol: 20 Issue: 02, October 2014.

[13] A Olson, D., *Data Mining Mo*dels, Business Expert Press 2016

[14] Ian H Witten; Eibe Frank; Mark A Hall; Christopher J Pal, *Data Mining : Practical Machine Learning Tools and Techniques, Fourth Edition*, Elsevier Science and Technology Books, Inc., 2017.

[15] http://www.meb.unibonn.de/epileptologie/science/physik/eegdata.html, Accessed 30 January 2018

[16] https://www.cs.waikato.ac.nz/ml/weka/, Accessed 30 January 2018

[17] V.P. Nigam and D. Graupe, "A neural-network-based detection of epilepsy," Neurology Research, vol 26, issue 6, pp 55–60, 2004.

[18] N. Kannathal, M.L. Choo, U.R. Acharya, and P.K. Sadasivan, "Entropies for detection of epilepsy in EEG," Computer Methods and Programs in Biomedicine, vol 80, issue 3, pp 187–194, 2005.

[19] A. Subasi, "EEG Signal classification using wavelet feature extraction and a mixture of expert model," Expert Systems with Applications, vol 32, issue 4, pp 1084–1093, May 2007.

[20] V. Srinivasan, C. Eswaran, and N. Sriraam, "Approximate entropy based epileptic EEG detection using artificial neural networks," IEEE Tran. Information Technology in Biomedicine, vol 11, issue 3, pp 288–295, May 2007.

[21] Hassan Ocak, "Automatic detection of epileptic seizures in EEG using discrete wavelet transform and approximate entropy," Expert Systems with Applications, vol 36, issue 2, pp 2027–2036, March 2009.

[22] R. Dhiman, J.S. Saini, and Priyanka, "Genetic algorithms tuned expert model for detection of epileptic seizures from EEG signatures," Applied Soft Computing, vol 19, pp 8-17, June 2014