MACHINE LEARNING BASED EEG SIGNAL CLASSIFICATION

Anjali Vasant Kadwe
Department of Electronics and Telecommunication
KCT's Late G. N. Sapkal College of Engineering,
Anjaneri, Nashik 422212

ABSTRACT

Epilepsy is a neurological disorder which is characterized by transient and unexpected electrical disturbance of the brain. The electroencephalogram (EEG) is a commonly used signal for detection of epileptic seizures. The proposed method is based on the classification of EEG signal with the less number of sample and more accurately by using the Matlab software. The Bonn University Data set use in this project provide classification of EEG signal by using the latest transform method. The project consist of Extraction of the data from text file,Frequency domain low pass filtering And Feature extraction by three most recent transform such as Coiflet Transform, Stationary Wavelet Transform (SWT) and Walsh Hadamard Transform (WHT). This transformed signal is the classified by KNN ensemble classification.This project provide an overall classification accuracy of 99%.

Keyword : - *Epilepcy, Electroencephalogram, Proposed work, feature extraction, Classification*

I. INTRODUCTION

Epilepsy is the brain disorder disease in which the neurons behave abnormal activity and this will lead to the seizure formation. Electroencephalogram (EEG) has played a vital role to monitor the activity of the brain ,especially in diagnosis of epilepsy. This disease is mostly found in adults at their old ages between 65 to 70 years or in kids and effects around 2% of the world population. Patients can be cured from this disease if they treated at its initial stage. If epileptic activity is recognized and operated appropriately, then 80% of the patients can be treated successfully. Due to the lack of state-of-the-art equipment and diagnosis tools in developing countries, the treatment of epileptic patients is not carried out properly. Normally, recording of EEG brain signals were examined by the neurophysiologist visually for identifying the epileptic seizure abnormalities in the signals. It is very time consuming and costly to inspect the epileptic signals visually. Moreover, human error may be involved in falsely detection and diagnosis of epilepsy which may cause harm to human life. Therefore, nowadays there is a need to develop an intelligent system to detect the epilepsy correctly, which will be helpful for timely and proper treatment of affected patients.

Different methodologies were utilized for epilepsy detection in the EEG signals such as artificial neural network, time-frequency component analysis, correlation function, time-domain analysis, frequency domain analysis and fuzzy logic based analysis. There are two steps involved in detecting the epilepsy in time-domain analysis, i.e., (1) extraction of discriminatory features from EEG brain signals from non-epileptic and epileptic subjects and (2) an intelligent pattern recognition system is formed for locating the abnormal activity (epileptic) in EEG signal within less time and high accuracy. Then decomposition of extracted features was carried out at every node of the tree. But we know that the EEG signals are non-stationary. Therefore, the best option to extract the suitable features from EEG signals is to use the time frequency based technique, i.e., DWT which is used to obtain the frequency and time information from the signals concurrently.

We decomposed the EEG epochs into different frequency bands up to 5th level by utilizing the db4 mother wavelet. Then we extract statistical wavelet features, i.e., energy, standard deviation and entropy from the wavelet bands. Finally, we use the state-of-the-art

machine learning technique, i.e., Support Vector Machines (SVM) for the training and classification of epileptic and non-epileptic signals on the extracted features. Our main contribution in this research paper is to analyze the recorded EEG signals using digital signal processing (DSP) tools like DWT and then classify them into various classes and extract the statistical wavelet feature. After extraction of discriminatory feature, we classify the EEG signals accurately. The prediction accuracy of the proposed pattern recognition system is 100% which confirms the significance of this system as compared to previous methods.

II. LITERATURE REVIEW

The recognition of epileptic and non-epileptic EEG signals is a classification problem. It involves extraction of the discriminatory features from EEG signals and then performing classification. In the following paragraphs, we give an overview of the related state-of-the-art techniques, which use different feature extraction and classification methods for classification of epileptic and non-epileptic EEG signals.

Yash paul et.al.[1] in 2018 worked on “Various epileptic seizure detection techniques using biomedical signals”. In this paper, seizure detection techniques are classified as time, frequency, wavelet (time–frequency), empirical mode decomposition and rational function techniques. The aim of this review paper is to present state-of-the-art methods and ideas that will lead to valid future research direction in the field of seizure detection.

Ihsan Ullah1 et.al.[2] in 2018 worked on “An Automated System for Epilepsy Detection using EEG Brain Signals based on Deep Learning Approach”.Epilepsy is a neurological disorder and for its detection, encephalography (EEG) is a commonly used clinical approach. Manual inspection of EEG brain signals is a time-consuming and laborious process, which puts heavy burden on neurologists and affects their performance. Several automatic techniques have been proposed using traditional approaches to assist neurologists in detecting binary epilepsy scenarios e.g. seizure vs. non-seizure or normal vs. ictal.

Shih-Kai Lin et.al.[3] in 2018 worked on “An Ultra-Low Power Smart Headband for Real-Time Epileptic Seizure Detection”. In this paper, the design of a smart headband for epileptic seizure detection is presented. The proposed headband consists of four key components: 1) an analog front-end circuitry; 2) an epileptic seizure detection tag (ESDT); 3) a Bluetooth low-power chip; and 4) customized electrodes. All the above components are integrated into a fabric headband with only 50.3 g.

Mohammad-Parsa Hosseini et.al.[4] in 2017 worked on “Random ensemble learning for EEG classification”. In this project a new method of feature selection and classifification for rapid and precise seizure detection is discussed wherein informative components of electroencephalogram (EEG)-derived data are extracted and an automatic method is presented using infinite independent component analysis (I-ICA) to select independent features. The feature space is divided into subspaces via random selection and multichannel support vector machines (SVMs) are used to classify these subspaces. The result of each classifier is then combined by majority voting to establish the final output.

III. PROPOSED WORK

Figure 3.1 shows the block diagram of complete system. The project consist of Extraction of the data from text file, Frequency domain low pass filtering And Feature extraction by three most recent transform such as Coiflet Transform, Stationary Wavelet Transform (SWT) and Walsh Hadamard Transform (WHT).

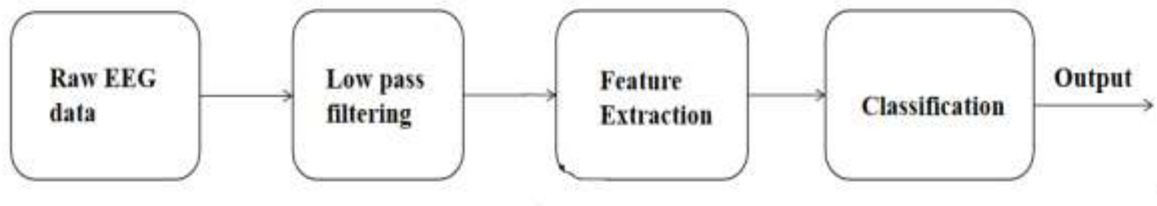

Figure 3.1 block diagram of the system

3.2 Raw EEG Data:

Extraction of the data from text file:

In this project we have used the A and E dataset from the Bonn University for detection and classification of the EEG signal and all this project is carried out with the help of MATLAB software so extraction of the data from the text file is the first step in the process as all the flow work of the process is shown in the fig 1 as follows:

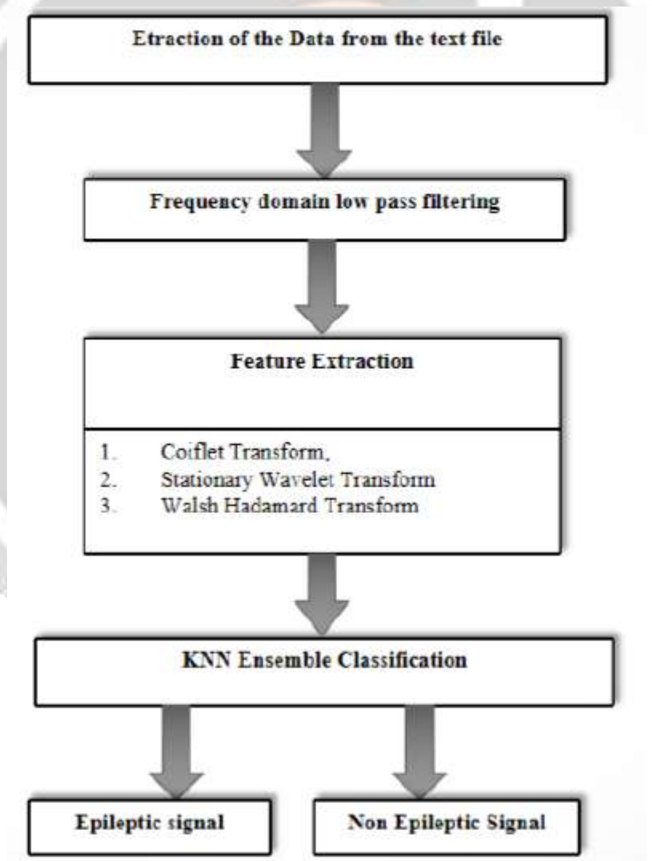

Figure 3.2 Work flow of the system

3.3 Preprocessing:

Frequency domain low pass filtering:

The frequency domain processing techniques are based on modifying the Fourier transform of an image. The basic idea in using this technique is to enhance the image by manipulating the transform coefficient of the image, such as Discrete Fourier Transform (DFT), Discrete Wavelet Transform (DWT), and Discrete Cosine Transform (DCT). This methods advantages includes low complexity of computations, ease of viewing and manipulating the frequency composition of the image and the easy applicability of special transformed domain properties.

An ideal low pass filter deals with the removal of all high frequency values of the Fourier transform that are at a distance greater than a specified distance from the origin of the transformed image. The filter transfer function for the Ideal low-pass filter is given by:

$H(u,v) = 1$ if $D(u,v) < D0$

$H(u,v) = 0$ if $D(u,v) > D$

3.4 Feature Extraction:

Feature extraction is either time domain or frequency domain for EEG signal. In most of the cases we have used different features in bio potential signals, this is because of the characteristics of the EEG signal. However, there are different frequency band of EEG signals, such as alpha beta, delta, and gamma. We have work out of our EEG dataset based on time domain and frequency domain features extraction directly from the signal. The important time domain features such as, maximum value, mean value, standard deviation, skewness, kurtosis etc., were extracted from raw EEG datasets. These are described below:

3.4.1 Mean Value :

The range of EEG potential is microvolts with time. The mean of EEG is constant and small values which change potentially. Definition of mean as described based on EEG signal is given by:

$$\mu_x = \frac{1}{N}\sum_{i=1}^{N} x_i$$

Where *x* is EEG signal and N is total number EEG data.

3.4.2 Standard deviation:

Standard deviation is a measurement to find the amount of variation or dispersion from the average. To determine standard deviation of EEG signal, we have considered low and high variation. The standard deviation can be expressed as

$$std = \sqrt{\frac{\sum(x - \mu_x)^2}{N - 1}}$$

3.4.3 Skewness*:*

It is a measurement of the asymmetry of the probability distribution of an EEG signal and it’s mean. The mean value can be positive, negative or undefined. The mean values can be zero based on completely symmetrical distribution and some nonzero values are asymmetrically distribution in respect of the baseline.

$$Skew = \frac{\sum_{i=1}^{N}(x_i - \mu_x)^3}{(n-1)\sigma_x^3}$$

3.4.5 Kurtosis:

It’s the measurement of the peakness of the probability distribution of EEG signal. The signal contains transient spikes, isolated high-voltage wave group. The kurtosis of it presents high positive values or negative values. These values are observed when EEG with high or low frequency and amplitude modulation is analyzed. The moment co-efficient can be written as

$$Kurt = \frac{\sum_{i=1}^{N}(x_i - \mu_x)^4}{(n-1)\sigma_x^4}$$

3.4.6 Coiflet Transform:

In the proposed method, the Coiflet wavelet was selected as the wavelet basis function. This wavelet exists under the name of Coiflets, but it is indeed constructed by I. Daubechies at the request of R. Coifman. Therefore, although Coiflet wavelet and well-known Daubechies wavelet are similar in a certain level, the Coiflet wavelet was indeed different in that it was constructed with vanishing moments not only for wavelet function Ã(t), but also for scaling function Á(t). In this way of design, the scaling function of the Coiflet will exhibit interpolating characteristics, which also implies that this wavelet allows a very good approximation of polynomial function at different resolution.

In the Coiflet wavelet formulation process, the following equations must be satisfied

$$\int dx\, x^l \psi(x) = 0 \qquad \text{for} \quad l = 0, \ldots, L-1$$

$$\int dx\, \phi(x) = 1$$

$$\int dx\, x^l \phi(x) = 0 \qquad \text{for} \quad l = 1, \ldots, L-1$$

Figs. 1(a) and (b) plot the corresponding scaling function and wavelet basis function. These functions were shown to be smoother and more symmetric than Daubechies wavelet [. This observation indicates that at different resolution, the approximation of polynomial functions can be better achieved. Furthermore, the symmetry property of the Coiflet is desirable in the signal analysis work due to the linear phase of the transfer function. As for the comparison with Morlet wavelet that was also tested in our laboratory [16—18], although the Coiflet method is less flexible in visualizing any frequency of interest, its discrete form is useful for the digital implementation. These benefits consolidate the utilization of Coiflet wavelet transform for this study.

3.4.6. Stationary Wavelet Transform (SWT):

The DWT suffers from time variant property. This means that the DWT of a translated version of a signal X is not the translated version of the DWT of X.

A. ε-decimated DWT :

There exist a lot of slightly different ways to handle the DWT. The decimation in DWT retrains even indexed elements, which is where the time variant problem lies in. The decimation could be carried out by choosing odd indexed elements instead of even indexed elements. The choice of even or odd concerns every step of the decomposition process. If we perform all the different possible decompositions of the original signal for a given maximum level J, then we will have 2J different decompositions [11].

Suppose εj=1 or 0 denotes the choices of odd or even indexed elements at step j. Then, every decomposition is labeled by a sequence of 0s and 1s, namely, ε=ε1ε2…εJ. This transform is called the ε-decimated DWT. A graphical example of ε=10110 is shown in figure

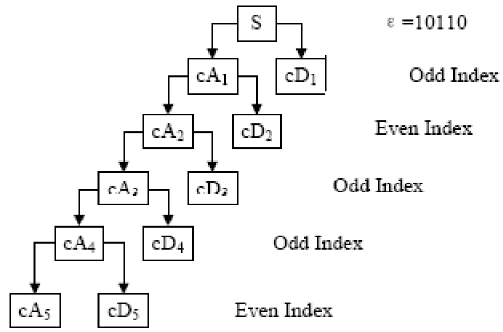

Figure 3.3 show graphical illustration of ε-decimated DWT

B.1D SWT :

The SWT can calculate all the ε-decimated DWT for a given signal at one time. More precisely, for level 1, the SWT can be obtained by convolving the signal with the appropriate filters as in the DWT but without downsampling. Then the coefficients of the approximation and detail at level 1 are the same as the signal lengthThe general step j convolves the approximation coefficients at level j-1, with appropriate filters but without down sampling, to produce the approximation and detail coefficients at level j. The schematic diagram is shown in Fig. 5.

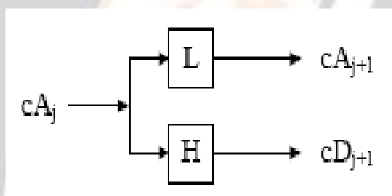

Figure 3.4 schematic diagram of 1D SWT

B. 2D SWT :

The algorithm of 1D SWT can be easily extended to the 2D case. Fig. 6 shows the schematic diagram of 2D SWT.

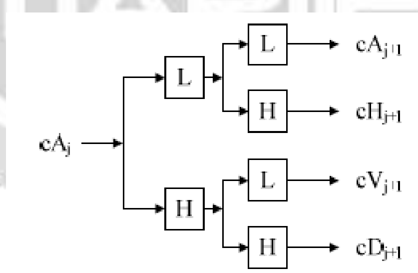

Figure 3.5 schematic diagram of 2D SWT

3.4.7 Walsh Hadamard Transform (WHT):

Linear image transforms [25] such as: Discrete Cosine Transform (DCT), Discrete Fourier Transform (DFT), Walsh Hadamard Transform (WHT) and Karhunen-Loeve transform (KLT) have been used in various image/video processing applications due to their energy compaction, entropy, flexibility, robustness and performance. Through analyzing or classifying the image frequencies or coefficients of the linear image transforms, edges of the image can be detected. The kernels or basis of the image transforms used for extracting edges can also be used to compute a set of energy measures which could characterize the local texture properties of a given region in an image. Hence, the kernels can extract the local texture property.

Among the linear image transforms, WHT is very attractive due to the simplicity of its implementation and to its properties which are similar to other transforms. Also, the computational efficiency of WHT makes it very attractive image processing directly in the transform domain since the components of the basis vectors are orthogonal and have only binary values (−1 or +1). WHT is a suboptimal, non-sinusoidal, orthogonal transform and it is used in many different applications, such as filtering, processing speech and medical signals, etc. More specifically, this transformation is used for astronomical signal/image processing, coding and filtering operations. WHT is well known for its simple and fast transformation.

The WHT is represented as a matrix and constructed from the WHT Matrix (WHTM) . An WHTM is defined as a set of N rows, denoted Wj , for j = 0, 1,…, N − 1, which have the following properties i) Wj takes on the values +1 and −1, ii) Wj[0] = 1 for all j, iii) Wj has exactly j zero crossings, for j = 0, 1,…, N−1. The size of a transform matrix is generally a power of 2. The matrix exists when N>2 and (N mod 4) = 0. The sequence ordered WHTM of order 4 is

$$For\ N = 4: \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \\ 1 & -1 & 1 & -1 \end{pmatrix}$$

Each row of WHTM is called a 1D Walsh Hadamard Basis vector. In general, basis vectors are orthogonal (dot product of any two of them is zero) and orthonormal (dot product of each with itself is 1). Using the tensor of 1-D basis vectors of WHTM, the 2D WHT kernels (WHTK) or basis images are generated by multiplying the corresponding row and columns as given below:

$$Tensor\ product\ of\ (2,2)\ \begin{array}{c c} & \begin{array}{cccc} 1 & 1 & -1 & -1 \end{array} \\ \begin{array}{c} 1 \\ 1 \\ -1 \\ -1 \end{array} & \begin{pmatrix} 1 & 1 & -1 & -1 \\ 1 & 1 & -1 & -1 \\ -1 & -1 & 1 & 1 \\ -1 & -1 & 1 & 1 \end{pmatrix} \end{array}$$

The above matrix is the tensor product of the 2nd row and 2nd column basis vector of WHTM. Each tensor product is a Basis image or Kernel of the WHT. Given an N × N block of pixels (where N must be a power of 2, N = 2n), its WHT is obtained by projecting the block over the WHT Kernels. However, these kernels can be represented as vectors forming the Basis Vectors of WHT and are represented in vector space as V = v1, v2,...,v16 named from left to right of the WHT Kernel given in Fig. 1. The WHT kernels (or basis images) for all basis vector of WHTM of order 4 is shown in Fig. 1 and the kernels below the diagonal representation are transpose of the kernels above the diagonal. From here onwards, the term basis vector represents the basis vector of WHT.

3.5 KNN ensemble classification:

3.5.1. Ensemble learning

Ensemble learning is an effective and increasingly adopted technique that combines multiple learning algorithms to improve overall performance. Ensemble learning consists of the following four main parts.

- Bagging: bagging, based on the majority-voting concept, utilizes randomly selected training data subsets, for training a dissimilar base learner of a similar manner.

- AdaBoost: adaboost is a famous member of the boosting approach. It creates base classifiers via sequential bootstrap samples, gained by weighting the training transactions via numerous iterations. Weighting is adjusted through mis classification related to the base classifier.

• Stacking: stacking combines different learning algorithms, in order to achieve higher prediction accuracy.

3.5.2 K-nearest neighbor

K-nearest neighbor (k-NN) is a simple classification model that exploits lazy learning. It is a supervised learning algorithm, which classifies new instance queries based on the majority of the k-nearest neighbor category. Calculating the minimum distance between the query instance and the training sample approximates the k-NN category. The k-NN prediction of the query instance is determined based on majority voting of the nearest neighbor categore

K-nearest neighbor (KNN) classifier works in a simple way by comparing both test and training data based on its nearest values [29]. In this case, „k‟ here refers to how many nearest value should be considered before the output class is decided. For instance, if k = 3, three nearest points between training and test data will be considered, and its final output belongs to the training data‟s class with most nearest points [30]. Implementation of KNN was accomplished using "Classification KNN" class in MATLAB‟s statistic toolbox. Various settings were given to customize the classification behavior of KNN, which include its output decision given that the same amount of closest point to multiple classes happened. In our case, the classifier uses random tiebreaker to assign the input data to a tied group randomly with the use of Euclidean distance as distance metric to compute data points‟ nearest neighbor. The said configuration has proven to provide high accuracy in classifying human stress [29]. Apart from its distance metric and tiebreaker rules, numbers of nearest neighbors to consider during classification is another vital setting. Through cross-validation, k-value of 2 is found to produce lowest average absolute error.

IV. EXPERIMENT AND RESULT

To examine the effectiveness of the proposed method and to prove its advantage over the other methods, the proposed method is experimented over the Bonn University dataset and it is evaluated by performance criteria. In this section, we focus mainly on dataset, performance measures and comparison of the proposed method with the existing methods.

4.1 Dataset and Evaluation Criterion:

To enable comparison with other other techniques, the proposed method was tested on the Matlab2018b software evaluation data. There are both epileptic and non epileptic sequences. n addition to this, in order to compare the proposed method with recent related works, the proposed method is tested with some software listed in Table I as follows:

4.2 Result Explanation:

Our algorithm validate the model by using 50% of the total training dataset and 50% of the testing dataset in model. Total number of sample of all dataset for training was 800 and testing 800. From these samples evaluation was take place and test accuracy evaluated. The performance (accuracy) of the proposed model with respect to training and testing is summarized in the following table:

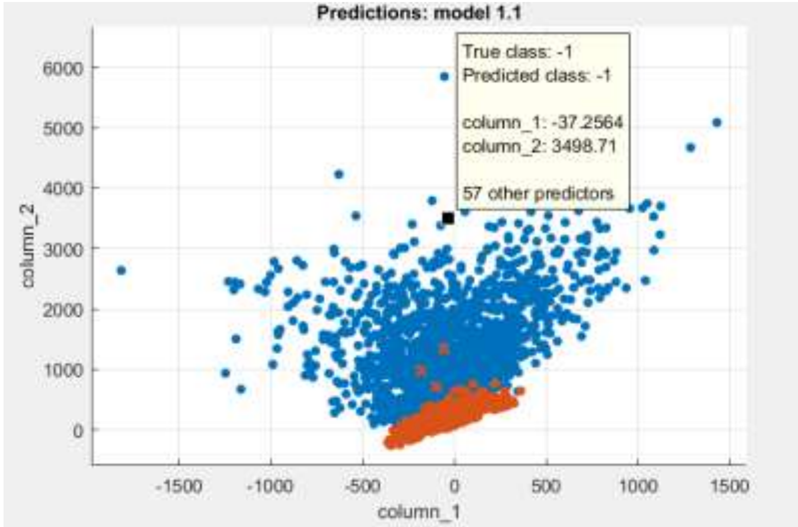

Figure 4.1 : Scattered plot of the result

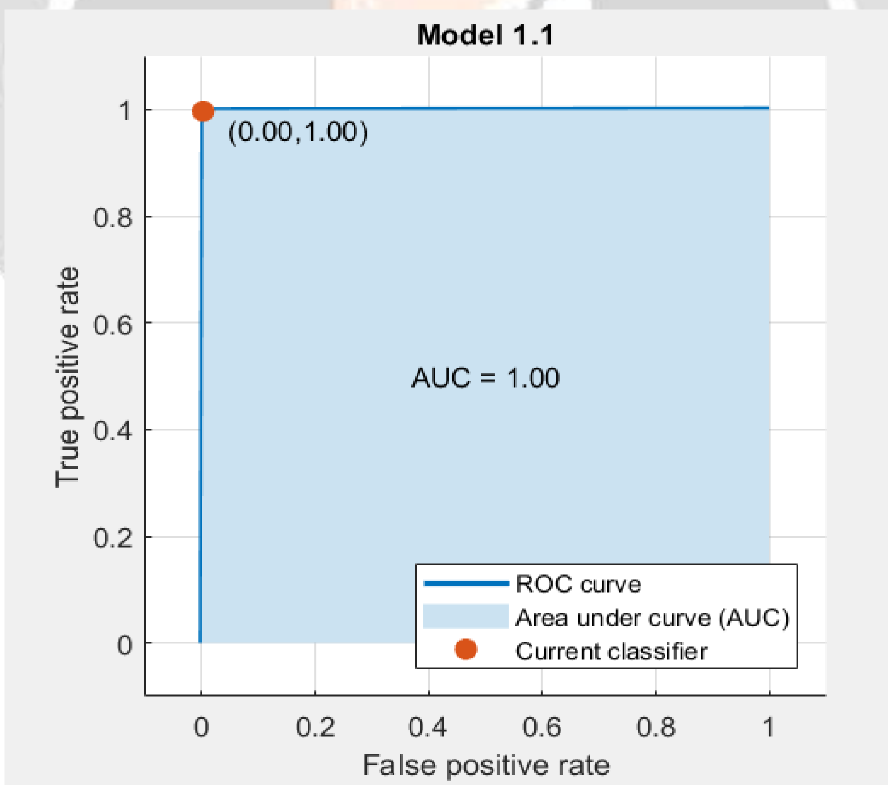

Figure 5.2 ROC curve of system

Sr.No.	Classifiers Name	Accuracy (%)
1	Medium Gaussian SVM	99.8
2	Coarse Gaussian SVM	99.1
3	Fine KNN	99.4
4	Medium KNN	98.7
5	Coarse KNN	94.9
6	Cosine KNN	95.3
7	Cubic KNN	97.0
8	Weighted KNN	98.7
9	Fine tree	99.8
10	Medium tree	99.8
11	Coarse tree	99.8
12	Logistic regression	99.8
13	Linear SVM	**99.9**
14	Cubic SVM	99.4
15	Quadratic SVM	99.8
16	Fine Gaussian SVM	99.0
17	Ensemble boosted tree	50
18	Ensemble bagged tree	99.8
19	Ensemble subspace discriminant	90.9
20	Ensemble subspace KNN	**99.9**
21	Ensemble RUS boosted tree	50

Table 4.1:Comparison results of existing feature extraction,classifications and performance techniques for EEG dataset.

Table 4.2 :Comparative Table of Accuracy, sensitivity, specificity with different Classification method

Sr.No.	Method	Accuracy	Sensitivity	Specificity
1	Ran Sub com of Class	0.97	0.98	0.96
2	Ran Sub NL SVM	0.95	0.96	0.94
3	Ran Sub L SVM	0.91	0.92	0.90
4	Non-linear SVM	0.85	0.85	0.87
5	ENN	0.85	0.84	0.83
6	Linear SVM	0.84	0.83	0.85
7	MLP neural network	0.82	0.81	0.83
8	**Ensemble KNN**	**0.9988**	**0.9988**	**0.9994**

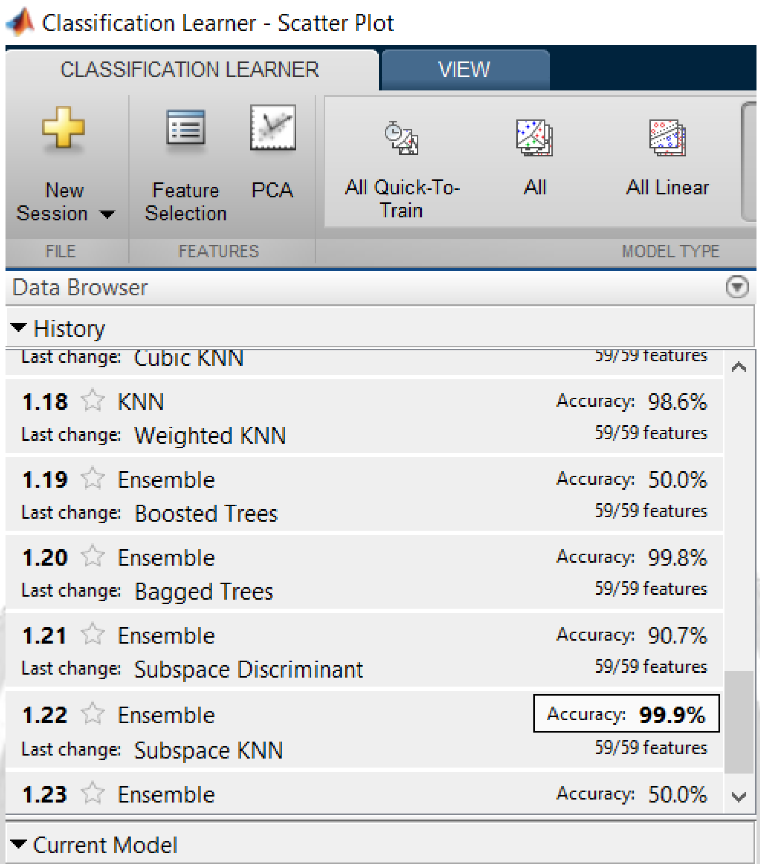

Figure 4.4 showing the output Accuracy as 99.9% of KNN classifier

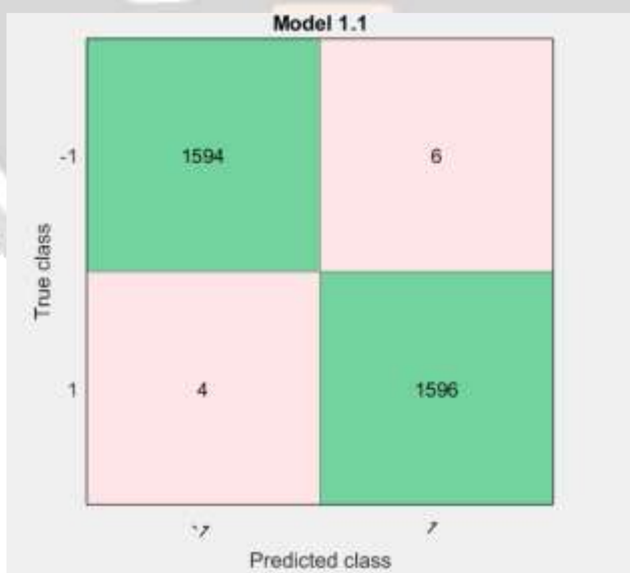

Figure 4.5 showing the confusion matrix of the process

V. CONCLUSIONS

This paper presented an EEG data classification algorithm to epileptic and non epileptic by using existing dataset which, based on a large number of features extracted after wavelet transform.We have studied the existing literature and found that very few approaches can successfully classify the EEG signal with less number of samples.It is challenging problem to be solved using machine learning techniques in coming months an efficient algorithm for EEG signal classification can be developed. Therefore, the conclusion is that the proposed algorithm can be used to classify EEG signals with 99% accuracy.

REFERENCE

[1] Yash paul "Various epileptic seizure detection techniques using biomedical signals" *springer.*

[2] Ihsan Ullaha Muhammad Hussainb Emad-ul-HaqQazib HatimAboalsamhb "An Automated System for Epilepsy Detection using EEG Brain Signals based on Deep Learning Approach" *Elsevier* https://doi.org/10.1016/j.eswa.2018.04.021

[3] Shih-Kai Lin, Istiqomah, Li-Chun Wang, Chin-Yew Lin, and Herming Chiueh"An Ultra-Low Power Smart Headband for Real-Time Epileptic Seizure Detection" *IEEE journal of Translational Engineering of Health and Medicine* 10.1109/JTEHM.2018.2861882

[4] Mohammad-Parsa Hosseiniae Dario Pompilia Kost Elisevichbc Hamid Soltanian-Zadehde "Random ensemble learning for EEG classification" *Artificial Intelligence in Medicine* https://doi.org/10.1016/j.artmed.2017.12.004

[5] Vipin Gupta Abhijit Bhattacharyya Ram Bilas Pachori "Classification of seizure and non-seizure EEG signals based on EMD-TQWT method" *2017 22nd International Conference on Digital Signal Processing* 10.1109/ICDSP.2017.8096036

[6] Evangelia Pippa, Vasileios G. Kanas, Evangelia I. Zacharaki, Vasiliki Tsirka, Michael Koutroumanidis, Vasileios Megalooikonomou "EEG-based Classification of Epileptic and Non-epileptic Events 2 using Multi-array Decomposition" *International Journal of Monitoring and Surveillance Technologies Research, 2017* HAL Id: hal-01359125

[7] Amirmasoud Ahmadi Vahid Shalchyan Mohammad Reza Daliri "A new method for epileptic seizure classification in EEG using adapted wavelet packets" *IEEE* 2017 Electric Electronics, Computer Science, Biomedical Engineerings' Meeting

[8] Lina Wang ,Weining Xue ,Yang Li ,Meilin Luo ,Jie Huang ,Weigang Cui [3] andChao Huang [3] "Automatic Epileptic Seizure Detection in EEG Signals Using Multi-Domain" *Information Theory Applied to Physiological Signals* https://doi.org/10.3390/e19060222

[9] Larbi Boubchir , Somaya Al-Maadeed , Ahmed Bouridane and Arab Ali Cherif"classification of EEG signals for detection of epileptic seizure activities based on LBP descriptor of time-frequency images"*IEEE International Conference* 978-1-4799-8339-1/15/

[10] Emad-Ul-Haq Qazi Muhammad Hussain Hatim Aboalsamh Wadood Abdul Saeed Bamatraf Ihsan Ullah."An Intelligent System to Classify Epileptic and Non-Epileptic EEG Signals" *12th International Conference on Signal-Image Technology & Internet-Based Systems* 10.1109/SITIS.2016.44

[11] Paul Fergus, David Hignett, Abir Hussain, Dhiya Al-Jumeily, and Khaled Abdel-Aziz "Automatic Epileptic Seizure Detection Using Scalp EEG and Advanced Artificial Intelligence Techniques"*BioMed Research International* http://dx.doi.org/10.1155/2015/986736

[12] Dragoljub Gajic, Zeljko Djurovic, Stefano Di Gennaro and Fredrik Gustafsson"Classification of EEG Signals for Detection of Epileptic Seizures Based on Wavelets and Statistical Pattern Recognition" *world scientific* https://doi.org/10.4015/S1016237214500215.