SIGNAL MODELLING APPROACH FOR CLASSIFICATION OF EPILEPTIC SEIZURE EEG SIGNAL USING SUPPORT VECTOR MACHINE

[1] Divya Gorde ,[2] Aashish Bardekar
[1] Student , [2] Professor
[1] Department of Computer Science and Engineering
[1] Sipna Collage of Engineering and Technology, SGBAU Amravati, India.

Abstract : This paper presents a signal modeling based new methodology of automatic seizure detection in EEG signals. A seizure is an event that causes an abrupt surge of electrical activity in the brain, where epilepsy is the disease involving recurrent unprovoked seizures. Worldwide, approximately 50 million people are affected by epilepsy and this is one of the most common neurological diseases. Electroencephalography (EEG) enlighten about the state of the brain i.e. about the electrical bustle going on in the brain. The electrical activity measured as voltage at different points of brain act as basis of EEG. These signals are generally time-varying and non-stationary in nature. Diagnosis of epilepsy requires long term electroencephalography (EEG) monitoring. The interpretation of long-term EEG monitoring takes a lot of time and requires the assistance of experienced experts. To optimize the seizure detection we used feature extraction concept. In order to overcome these limitations we apply machine learning technique.

***Index Terms* - Detection, EEG , Epilepsy, Machine Learning Technique, Seizure ,Non- Seizure, Signals.**

I. INTRODUCTION

Human brain is a highly complex system. The epilepsy is a common neurological disorder of human brain. It affects at least 50 million people of the world. The annual occurrence of epilepsy, 48 per 100,000 populations in developed countrie. The prevalence of epilepsy is higher in low and middle income countries than developed countries . At least 50 % of the epileptic cases start developing at childhood or adolescence. Electroencephalography (encephalon = brain), or EEG, is the physiological method of choice to record all of the electrical activity generated by the brain from electrodes placed on the scalp surface. For faster application, electrodes are mounted in elastic caps similar to bathing caps, ensuring that the data can be collected from identical scalp positions across all respondents. The diagnosis of epilepsy relies on the clinical history of the patient, computed tomography (CT), magnetic resonance imaging (MRI), video-recording, and electroencephalography (EEG). In particular, EEG is a common procedure used to support the diagnosis of epilepsy because it provides objective information. This information is in the form of a voltage-versus-time graph, recording the spontaneous electrical activity of the cerebral cortex.

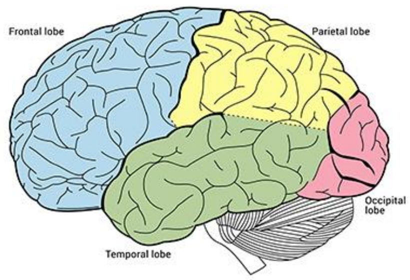


Fig. 1 Lobes of Brain

Epilepsy is a brain disorder in which clusters of nerve cells, or neurons in the brain sometimes have abnormal signal. In the epilepsy, the normal pattern of neurons activity becomes disturbed causing strange sensation, emotion, behavior and loss of consciousness. Epilepsy is a disorder due to many possible causes. Anything that disturbs the normal pattern of neuron activity may result in illness, brain damage, or abnormal brain development. EEG scan is a common diagnostic test for epilepsy and can detect abnormalities in the brain electrical activity. EEG plays an important role in the diagnosis of epilepsy. An EEG technically places the electrodes in specific areas, according to internationally agreed-upon criteria. The names of the electrode sites use alphabetical abbreviations that identify the lobe or area of the brain which each electrode records from: F = frontal ,Fp = frontopolar ,T = temporal ,C = central , P = parietal , O = occipital ,A = auricular (ear electrode) .The localization of the brain waves within the brain regions or lobes is further narrowed by adding electrodes, which are given numbers such as T3, T4, P3, P4. Even numbers identify electrode positions on the right side of the head, and odd numbers refer to the left side. The label "z" points to electrode sites in the midline of the head. For example, Cz refers to the midline central region of the head.

II. LITERATURE REVIEW

Epilepsy is the 4th most common neurological disease in the world. Worldwide, approximately 50 million of people are affected by Epilepsy and it is one of the most common neurological disease. We studied the techniques of seizure and non- seizure detection by different authors. They are given below:

Miran Lee, Inchan Youn, Jaehwan Ryu, Deok-Hwan Kim[1] was proposed the technique of a novel feature extraction method, a slope of counting wavelet coefficients over various thresholds (SCOT) method based hidden markov model (HMM) for seizure detection. The proposed SCOT method based HMM has a robust detection accuracy, and a short feature extraction time. Experimental result shows that with the proposed method, the average detection accuracies are 96.5% and 98.4% using the HMM in seizure and non-seizure, respectively. Finally, In order to solve the problem of long-term EEG monitoring, we calculated the feature calculation time of the SCOT method and compared it with the existing performance.We proved that SCOT is an appropriate feature method for real-time system and long-term monitoring.

J Satheesh Kumar and P Bhuvaneswari [2] focused on the EEG signals and its characteristics with respect to various states of the human body and also deals with experimental setup of EEG analysis. Electroencephalography (EEG) is an efficient modality which helps to acquire brain signal corresponds to various states from the scalp surface area. These signals are generally categorized as delta , theta, alpha , beta , and gamma based on signal frequencies ranes from 0.1Hz to more than 100Hz. The rapid advancements in biomedical technology for analysis bio medical signals are an important research area. One such technology is EEG , which is to measure the brain potential in order to help the disable people and obtain the accurate diagnosis of diseases. EEG records brain waves with respect to specific frequency by placing metal electrodes on the scalps.

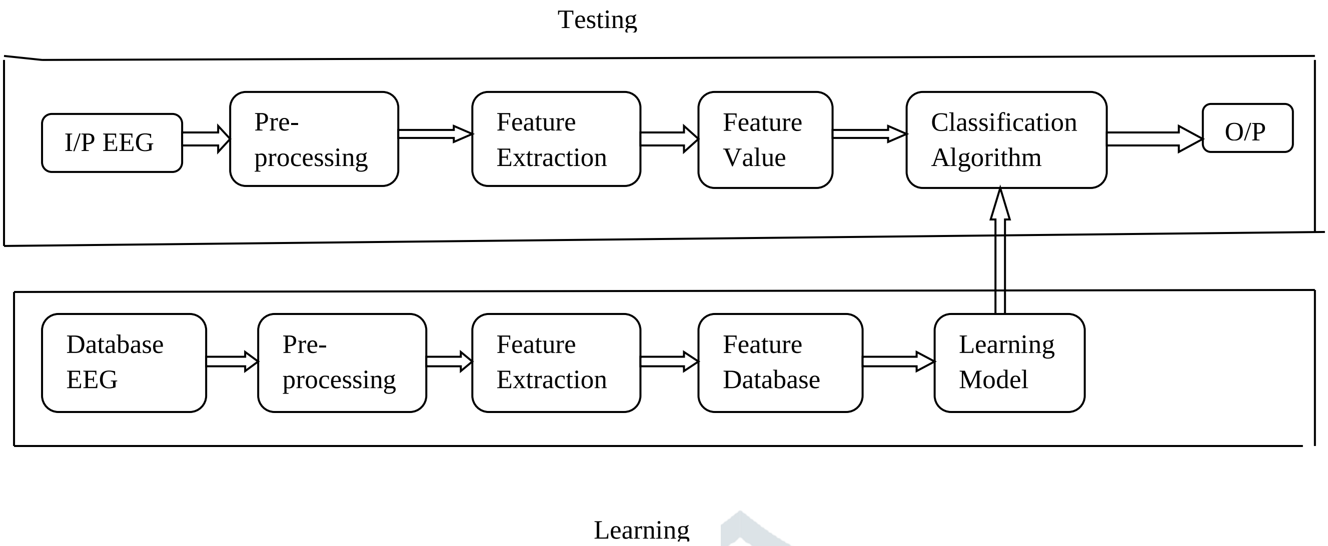
Tingxi Wena , Zhongnan Zhanga [3] told a genetic algorithm-based frequency-domain feature search (GAFDS) method is proposed for the electroencephalogram (EEG) analysis of epilepsy. In this method, frequency domain features are first searched and then combined with nonlinear features. Subsequently, these features are selected and optimized to classify EEG signals. The extracted features are analyzed experimentally.

Harender , R. K. Sharma [4] presents a framework for epileptic seizure detection from recorded EEG signal for healthy and epileptic patient. Simulink has been used to model, EEG signal decomposition using discrete wavelet transform (DWT) and statistical calculation; After DWT decomposition Mean Absolute Value (MA), Standard Deviation (SD) and Average power(AP) are extracted as statistical feature for epilepsy detection with k- Nearest Neighbor (k-NN) classifier. Results show that k-NN classifier gives better accuracy with SD and SD with MA for eyes open and epileptic seizure dataset with less number of extracted features.

Malik Anas Ahmad , Waqas Majeed ,et al [5] given the approach in diagnosis of epilepsy, computer aided methods can significantly supplement a neurologist by automatically identifying the epileptic patterns in an EEG. In the last decade immense amount of work has been done in the field of EEG based computer aided diagnosis of epilepsy. Computer assisted analysis of EEG for diagnosing Epilepsy significantly helps a neurologist. To avoid misinterpretation and over-interpretation a computer assisted system should be user friendly, accurate, robust and above informative.

III. PROPOSED WORK

In the proposed work , we have design the block diagram in that there are two phases. One is learning and other is testing. First the system will learn the data input and then test it in testing phase to generate the output.



1. Database :

The database which we are using is collected at the Children's Hospital Boston, consists of EEG recordings from pediatric subjects with intractable seizures. Subjects were monitored for up to several days following withdrawal of anti-seizure medication in order to characterize their seizures and assess their candidacy for surgical intervention. Recordings, grouped into 23 cases, were collected from 22 subjects (5 males, ages 3–22; and 17 females, ages 1.5–19). The file SUBJECT-INFO [SUBJECT-INFO] contains the gender and age of each subject. Each case (chb01, chb02, etc.) contains between 9 and 42 continuous .edf files from a single subject. All signals were sampled at 256 samples per second with 16-bit resolution. Most files contain 23 EEG signals (24 or 26 in a few cases). The International 10-20 system of EEG electrode positions and nomenclature was used for these recordings. In some cases, up to 5 "dummy" signals (named "-") were interspersed among the EEG signals to obtain an easy-to-read display format; these dummy signals can be ignored.

2. Preprocessing :

The electrodes are arranged according to 10-20 standards for EEG placement. These electrodes are label by letters (i.e. F-Frontal , T-Temporal , C-Central, P-Parietal, O-Occipital) which indicates the lobes of the brain. Midline region is label with 'z'. The position of EEG electrodes for recording EEG signals. According to the international 10-20 system standards, all 19 channels (FP1, FP2, F3, F4, F7, F8, T3, T4, T5, T6, O1, O2, P3, P4, C3, C4, FZ, CZ, PZ) and both ear lobes are used for the positions of the EEG electrodes for recording EEG signals.

3. Feature Extraction :

We have extracted features of the EEG signal of each lobe wise channels. It gives us the better accuracy for comparing the behavioural change in the EEG signal of epileptic patients and non-epileptic patients. The features we have extracted are given below:

- Mean: The mean of the data set is the arithmetic average of the elements in a data set obtained by adding all the values and dividing it by the number of values. In case of the data if in the form of frequency distribution, then the mean of frequency distribution data can be define as ,

$$\mu = \frac{1}{n}\sum_{i=1}^{n} f_i\, x_i = \sum_{i=1}^{n} p_i\, x_i$$

- Variance: The variance of data set is the arithmetic average of squared differences between the mean. Again, when we summarize a data set in frequency distribution, the variance of frequency distribution is given by,

$$\text{Variance} = \sigma^2 = \frac{1}{n}\sum_{i=1}^{n} f_i(x_i - \mu)$$

- Standard Deviation: The standard deviation(STD) of a data set in a frequency distribution can be define by the equation,

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

- Max: The highest possible value or the most significant value in a given data set is called maximum value (MAX).

- Kurtosis: It is define as the second measure of peaked distributions of potential values of activity values in each trial.

$$k = \frac{E(x-\mu)^4}{\sigma^4}$$

- Skew ness: It is defined as the measure of the similarity or asymmetric distribution. It can be positive as well as negative.

$$s = \frac{E(x-\mu)^3}{\sigma^3}$$

4. Classification Algorithm :

In machine learning, support vector machines SVMs, also support vector networks are supervised learning models with associated learning algorithms that analyze data used for classification and regression analysis. Given a set of training examples, each marked as belonging to one or the other of two categories, an SVM training algorithm builds a model that assigns new examples to one category or the other, making it a non-probabilistic binary linear classifier. An SVM model is a representation of the examples as points in space, mapped so that the examples of the separate categories are divided by a clear gap that is as wide as possible. In addition to performing linear classification, SVMs can efficiently perform a non-linear classification using what is called the kernel trick, implicitly mapping their inputs into high-dimensional feature spaces.

In the linear classifier model, we assumed that training examples plotted in space. These data points are expected to be separated by an apparent gap. It predicts a straight hyperplane dividing 2 classes. The primary focus while drawing the hyperplane is on maximizing the distance from hyperplane to the nearest data point of either class. The drawn hyperplane called as a maximum margin hyperplane.

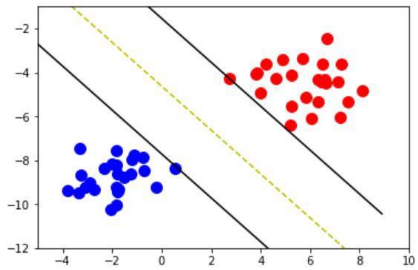
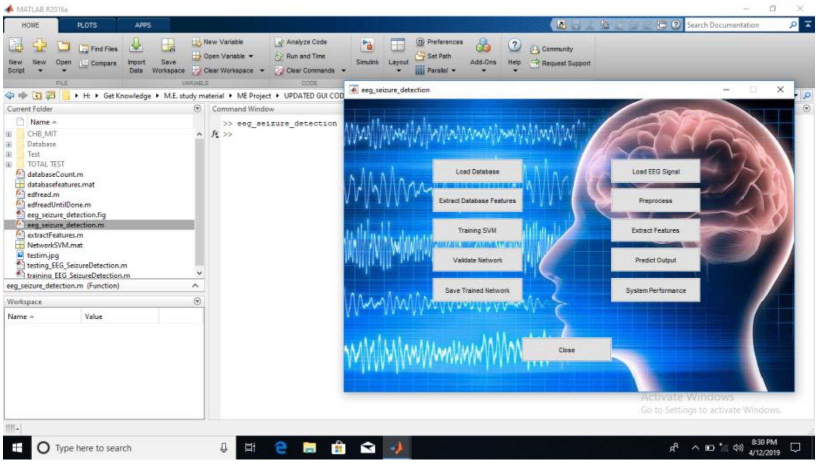


Fig.2 SVM with hyperplaine

In the field of machine learning and specifically the problem of statistical classification, a confusion matrix, also known as an error matrix, is a specific table layout that allows visualization of the performance of an algorithm, typically a supervised learning one (in unsupervised learning it is usually called a matching matrix). Each column of the matrix represents the instances in a predicted class while each row represents the instances in an actual class (or vice versa). It is a special kind of contingency table, with two dimensions ("actual" and "predicted"), and identical sets of "classes" in both dimensions (each combination of dimension and class is a variable in the contingency table).

IV. RESULT AND DISCUSSION

Here we use the software tool MATLAB. Matlab is a high-level technical computing language and interactive environment for algorithm development, data visualization, data analysis, and numeric computation developed by MathWorks. The version of matlab use is R2016a , 64-bit. The hardware requirement of system comes with Operating System is Microsoft Windows 10 Pro , System Type is x64-based PC , Processor is AMD A6-6310 APU with AMD Radeon R4 Graphics.The database which we are using is collected at the Children's Hospital Boston, consists of EEG recordings from pediatric subjects with intractable seizures. Below is the screenshot how our system looks in matlab software after design .Consider the data of any one patient give the input to the system and according to that it generate output which predict the seizure and non seizure signal.



Below is the table showing performance of system:

Parameter	Value
Accuracy	95.00%
True Positive	88.89%
False Positive	0.0%
True Negative	100%
False Negative	11.1%
Error Rate	5.00%
Precision	100%
Sensitivity	88.89%
Recall	88.89%
Specificity	100%
Force measure	97.56%

V. CONCLUSION AND FUTURE SCOPE

The EEG signals have gained a lot of importance in the field of biomedical science in the past few decades. The advancement in technology and its ever increasing demands have encouraged the engineers to ascertain new methods for analyzing these signals. Here we take support vector machines or the neural networks can be used for the classification of signal. EEG provides important information for epilepsy detection. Feature extraction, selection, and optimization methods exert significant influence in EEG classification. Our proposed system is to predict the seizure and non seizure EEG signal and by using statistical parameter gives more accurate result and less error rate.

VI. ACKNOWLEGMENT

A moment of pause, to express a deep gratitude to several individuals, without whom this paper could not have been completed. I feel immense pleasure to express deep sense of gratitude and indebtedness to my guide Prof. A.A.Bardekar for constant encouragement and noble guidance. I would like to thank Dr.S.A.Ladhake, Principal, Sipna college of Engineering andTechnology,Amravati for his kind co-operation, Guidance and for providing necessary facilities during the period of working on this paper. I express my sincere thanks to Dr.S.S.Dhande H.O.D Department of Computer Science and Engineering for their kind co-operation.

REFERENCES

[1] Miran Lee, Miran Lee, Inchan Youn, Jaehwan Ryu, Deok-Hwan Kim, “Classification of both seizure and non seizure EEG signal using Hidden Markov Model”, 2018 IEEE International Conference on Big Data and Smart Computing.

[2] J Satheesh Kumar and P Bhuvaneswari, "Analysis of Electroencephalography(EEG) Signals and Its Categorization – A Study", International Conference on Modeling , Optimization and Computing(ICMOU 2012).

[3] Tingxi Wena , Zhongnan Zhanga ,"Effective and Extensible Feature Extraction Method Using Genetic Algorithm-Based Frequency-Domain Feature Search for Epileptic EEG Multi-classification

[4] Harender , R. K. Sharma , "DWT based Epileptic Seizure Detection from EEG signal using k-NN classifier" ,International Conference on Trends in Electronics and Informatics ICEI 2017

[5] Malik Anas Ahmad, Waqas Majeed and Nadeem Ahmad Khan, "Advancements in Computer Aided Methods for EEG-based Epileptic Detection", BIOSIGNALS 2014 - International Conference on Bio-inspired Systems and Signal Processing.

[6] Anubha Gupta, Pushpendra Singh, Mandar Karlekar, "A Novel Signal Modeling Approach for Classification of Seizure and Seizure-free EEG Signals",2018 IEEE Transactions on Neural Systems and Rehabilitation Engineering .

[7] Ashley N. Johnson, Daby Sow, Alain Biem, "A Discriminative Approach to EEG Seizure Detection".

[8] Kartikeya Shrikant Bagalore, "Studying the Use of Hidden Markov Models in the Detection and Classification of EEG Epileptiform Transients using LPC features", (2016)

[9] Sabrina Ammar and Mohamed Senouci , " Seizure Detection with Single-Channel EEG using Extreme Learning Machine", 17th international conference on Sciences and Techniques of Automatic control & computer engineering - STA'2016, Sousse, Tunisia, December 19-21, 2016.

[10] Abdul Quaiyum Ansari , Priyanka Sharma , " A Review on Automated Detection of Nonconvulsive seizures using EEG", Second International Conference on Computational Intelligence & Communication Technology ,2016.

[11] Tarak Das ,Sayanti Guha, " Classification of EEG Signals for Prediction of Seizure using Multi-Feature Extraction" , IEEE 2017.

[12] www.epilepsysociety.org

[13] www.epilepsyontario.org

[14] www.physionet.org

[15] Md Kafiul Islam, Amir Rastegarnia, and Zhi Yang , "A Wavelet-Based Artifact Reduction from Scalp EEG for Epileptic Seizure Detection", IEEE Journal of Biomedical and Health Informatics, VOL. X, 2015.