



IJRASET

International Journal For Research in
Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 6

Issue: II

Month of publication: February 2018

DOI:

www.ijraset.com

Call:  08813907089

E-mail ID: ijraset@gmail.com

Classification of epileptic EEG signals based on J48 Classifier and Correlation based feature selection

Moushmi Kar¹, Laxmikant Dewangan²

^{1,2} Department of Electronics and Telecommunication, Choukseys Engineering College, Bilaspur (C.G.)

Abstract: *Epileptic seizures occurs when, there is an abrupt and huge surge of energy discharge by the brain cells which is uncontrolled. The EEG signals are using to determine seizure type and syndrome related with epilepsy in patients. It can helpful for multi axial diagnosis of epilepsy, in terms of whether the seizure disorder is focal or generalized, idiopathic or symptomatic or part of a particular epilepsy. This paper presents a new method which generate and selects features from EEG signals. In this paper, firstly, we used amplitude values of EEG data as features of epilepsy classification. Secondly, the correlation based feature selection (CFS) technique is applied to select the key features and to reduce the dimensionality of the data. Finally, the selected features are forwarded to J48 classifier to classify the EEG signals. The J48 classifier classified the features which are extracted from EEG signal and selected from the CFS. It is found that the proposed method achieves 94.357, 88.094 and 95.811, 91.952 % for classification accuracy, sensitivity and specificity, and area under curve respectively.*

Keywords: *Epilepsy, EEG, CFS, J48*

I. INTRODUCTION

Epilepsy is a chronic brain disorder, characterized by seizures, affecting approximately 70 million people around the world. It is characterized by recurrent convulsions over a time-period. The episodes may vary as low as once in a year to frequent fits occurring several times per day. When a sudden disturbance occurs in brain function, than it indicate the clinical signs of excessive and hyper-synchronous activity of neurons in the brain, is called seizures [29, 30]. The probable reason behind this behavioural disturbance is a kind of brief electrical "storm" arising from inherently unstable neurons due to presence of a genetic defect (as in the various types of inherited epilepsy). The other reasons by which neurons are being unstable are presence of abnormal metabolites such as low blood glucose, or alcohol. Sometimes this happens only occasionally; for others, it happens up to hundreds of times a day. Epilepsy is determined by EEG signal recording, which contain valuable information for understanding epilepsy. Epilepsy can create clear disturbance and leaves its signature on standard EEG signals. The brain activities are measured using noninvasively electroencephalography (EEGs) signals or invasively electrocortico graphy (ECoG). EEG and ECoG signals have been used to recognize, analyses and treat a number of neurological disorders and abnormalities of the brain. To estimate the significant information from EEG signals, many researchers developed new techniques. This information is used as the input to different classifier. There is various ways to extract the key features and to select the features. EEG signals are captured abnormal activities in any types of seizure. Seizure can cause permanent secondary damage during the initial stages of the epileptic seizure due to this it requires constant observation of patient by medical or nursing staff. this procedure is very laborious, costly, and not a practical solution. Automatic detection of seizure and spike methods can help the clinical staff as well as the EEGer to quickly review the prolonged EEG recordings to retain sections from recordings with clinical significance. It is found that there is a need for automated detection and classification of seizure.

One of the methods used in this paper for extracting epileptic EEG data is to choose amplitude of the EEG signal from EEG data set. In this technique, each amplitude of the data has the same chance to be selected as a subject. Then, we forwarded all these samples to the Correlation based feature selection (CFS) method for selecting the best features [76]. This study uses the selected features as the input for a classifier. One of the most popular classifiers, the decision tree classifier J48 is used to classify EEG data [83]. This technique is used to identify the epileptic and non epileptic EEG data.

Many researchers are developed different features extraction and feature selection method for detection of epileptic seizures. There are lots of approaches used by researchers to classify the epileptic seizure form EEG signals. Few discussion of previous study is provided below.

Kemal Polat et.al. Proposed a hybrid system based on Fast Fourier Transform and decision tree classifier [34]. In this study same dataset is used for classify the epilepsy. All data set were obtained from EEG signals of healthy subjects and subjects suffering from epilepsy diseases. In this study, he obtained 98.68% and 98.72% classification accuracies using 5- and 10-fold cross-validation.

Srinivasan et al. proposed an automated epileptic EEG detection system using approximate entropy as the feature in Elman and probabilistic neural networks. Elman network yielded an overall accuracy of 100% [36].

Ocak performed a method for automated seizure detection based on ApEn and DWT. The accuracy of seizure detection is more than 96% [37]. Acharya et al proposed four nonlinear parameters namely CD, FD, H and ApEn in SVM and GMM classifiers in the three class epilepsy detection [50]. It is found that, GMM classifier showed better performance with an average classification accuracy of 95%, sensitivity of 92.22%, and specificity of 100%. Musa et.al proposed a new method for the diagnosis of epilepsy from electroencephalography (EEG) signals based on complex classifiers [67]. It is reported that accuracy obtained from the above experiment is 98.28% for ten-fold cross validation and 100 % from hold out. Ashwani et. al proposed a method based on key point LBP for automated diagnosis of epilepsy from EEG signals [69]. The above mentioned studies have some drawbacks regarding the computational complexity. In this study, we are proposed a method which provides the best feature selection and classification with low computational complexity. The rest of the paper is organized as follows: Section 2

Explains proposed material and methodology Section 3 discusses the experimental results. Section 4 concludes the paper and mentions future work.

II. MATERIAL AND METHODOLOGY

A. Material

In this study, the publicly available EEG data from Bonn University is used for feature generation and selection. The complete data set includes five sets (denoted A–E) each containing 100 single channel EEG segments of 23.6 s. This study classify epileptic and non epileptic data set ,which selected the set A which was taken from surface EEG recordings of five healthy people with eye open, and set E which was taken from EEG records of five pre-surgical epileptic patients during epileptic seizure activity. In this section we explain only a short description and refer to Andrzejak et.al (2001) for further details [28].

B. Methodology

The research work proposed here is to generate the features of the EEG signals to identify the epileptic seizure and to classify the seizure activity accurately. The objective of this paper is to come up with best classifier with CFS feature selection technique. Finally, the classifier performance evaluated by cross validation partition method.A detail description of the proposed algorithms is discussed in the following section.

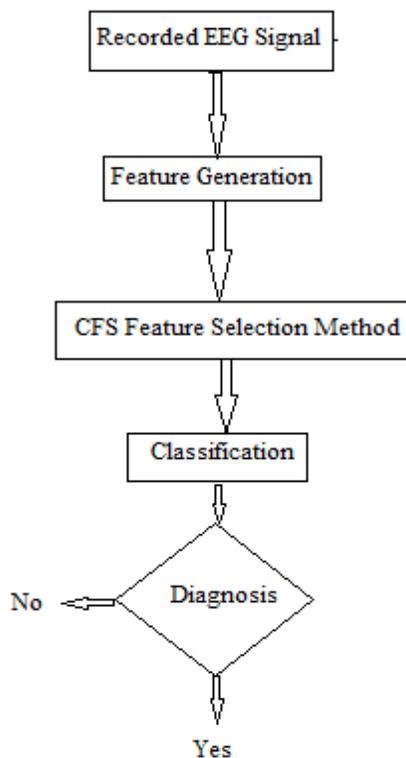


Fig.1 The flow chart of proposed methodology

This proposed method consists of three stages: feature generation , CFS feature selection and classification by J48 classifier In our method, first, we generate the 178 feature from EEG signals and then top 20 features are selected by CFS feature selection method and it was applied to decision tree J48 classifier to detection of epileptic seizure. The flow chart of the proposed method is given in Fig. 1.

- 1) *Feature Generation and Feature Selection:* The amplitude of recorded EEG data is applied as features vector of proposed method. Next, Features are applied to the correlation feature selection method. The process of feature selection is important because large set of features may contain irrelevant data. The challenge is to try to reduce the number of dimensions in which we are computing. This task is performed by feature selection method. In this paper, the correlation based feature selection method is applied for the next section. CFS evaluates the merit of feature subsets by using a search algorithm with a function. CFS provided an optimal solution by measure the goodness of feature subset taking into the account of the usefulness of individual features for predicting the class label along with the level of inter correlation among them [76].
- 2) *Classification:* Now the best top 20 feature selected is applied to the J48 classifier to classify the EEG signal into epileptic and non epileptic seizure. J48 classifier is a machine learning model which is predictive. This classifier is decided the dependent variable of a new sample based on various features values of the available data. The J48 decision tree classifier is depended upon the simple algorithm. In J48 classifier, classification is done by to create a decision tree based on the feature value of the available training data. In this method, the classification is performed by the WEKA software.
- 3) *Performance Measure:* In this section, the performance measurement of classifier is evaluated by different measure namely accuracy (also known as recognition rate), sensitivity (or recall) and specificity. The accuracy of a classifier is defined the correctly classify sample means, EEG samples are classified as epileptic and non epileptic samples. Sensitivity is also called true positive rate. It is defined as the percentage of seizure samples correctly classified as seizure. It is also called true negative rate. It is defined as the percentage of EEG samples correctly classified as epileptic.

III.RESULT AND DISCUSSION

In this Section, the proposed methodology is applied on the available database as discussed Section 2. The technique is employed to classify two-class EEG signals from datasets: epileptic and non epileptic data. After generate the feature vector the top 20 feature selected by CFS feature selection and classification done by J48 classifier. Table 1 explain the top 20 best feature vector selected by CFS feature selection method. Table 2 shows the performance measure of J48 classifier model using only top 20 relevant features for CFS feature selection technique.

TABLE I
TOP 20 FEATURES SELECTED BY CFS FEATURE SELECTION METHOD

Feature selection Method	Selected features
Correlation based feature selection (CFS)	X1,X3,X6,X8,X12,X16,X19,X24,X25,X30,X32,X35,X38,X41,X44,X48,X51,X56,X60,X64,

TABLE II
PERFORMANCE MEASURE OF DIFFERENT CLASSIFIER MODELS FOR TOP20 FEATURES USING 10 FOLD APPROACHES

Classifier Model	Performance Measure (%)				
	Accuracy	Sensitivity	Specificity	AUC	MCC
J48	94.357	88.094	95.811	91.952	0.820

IV. CONCLUSIONS AND FUTURE WORK

In the medical field research, the classification of EEG signal is a key issue for identification and evaluation of the brain activity. The EEG signals analysis require a large sets of EEG data and features used from this large data play an important role in

classifying EEG signals. In this dissertation, EEG signal processing and classification techniques in order to detect epilepsy is studied and developed. In this paper the J48 classifier classify the epileptic data by the accuracy of 94.357% and 88.094%, 95.811 %, 91.952 % and 0.820 sensitivity, specificity, AUC and MCC respectively. In this study, it is concluded that the proposed classifier such as decision tree J48 classifier can be useful for the epilepsy diagnosis.

A. Future Work

Future, assessment of proposed classifier in clinical practice and computer aided diagnosis on different dataset can be explored. Accurate classification of epileptic seizure and additional public facilities like computer aided diagnosis (CAD) can bring about the much needed improvement in epileptic seizure care.

REFERENCES

- [1] A. S., Zandi, M., Javidan, G. A Dumont, and R. Tafreshi, "Automated real-time epileptic seizure detection in scalp EEG recordings using an algorithm based on wavelet packet transform". IEEE Trans. Biomed. Eng., 57(7): 1639-1651, 2010.
- [2] B Yaffe , Philip Berger ,Pierre Megev and,Groppe,David M., Kramer Mark,A., Chu ,Catherine J., Santaniello Sabato., Meisel, Christian , Mehta, Ashesh D. , Sarma ,V.Sridevi. "Physiology of functional and effective networks in epilepsy". Clin. Neurophysiol., 126(2) 227-236, 2015.
- [3] M. Hall., "Correlation based feature selection for machine learning". Doctoral dissertation, University of Waikato, Dept. of Computer Science, 1999.
- [4] R. Quinlan, "C4.5: Programs for Machine Learning. Morgan Kaufmann Publishers", San Mateo, CA., 1993.
- [5] K Polat., and S. Günes., "Classification of epileptiform EEG using a hybrid system based on decision tree classifier and fast Fourier transform." Appl. Math. Comput, 187: 1017-1026, 2007.
- [6] V., Srinivasan, C. Eswaran, and N. Sriraam, "Approximate entropy-based epileptic EEG detection using artificial neural networks" IEEE Trans. Inform. Technol. Biomed. 11(3): 288-295, 2007.
- [7] H. OcaK, "Automatic detection of epileptic seizures in EEG using discrete wavelet transforms and approximates entropy" Expert Syst. Appl. 36(2): 2027-2036, 2009.
- [8] U. R. Acharya, K. C. Chua, T.C Lim, Dorothy, J.S. Suri, 'Automatic identification of epileptic EEG signals using nonlinear parameters'. J. Mech. Med. Biol. 9(4): 539-553, 2009.
- [9] M., Peker, B. Sen, and D. Delen, "A Novel Method for Automated Diagnosis of Epilepsy Using Complex-Valued Classifiers" IEEE Journal of Biomedical And Health Informatics. 20(1), 2016.
- [10] A. K.Tiwari, R. B. Pachori, V. Kanhangad, and B. K. Panigrahi, Automated Diagnosis of Epilepsy Using Key-Point-Based Local Binary Pattern of EEG Signals. IEEE Journal of Biomedical And Health Informatics. 21(4), 2017.
- [11] RG Andrzejak, K Lehnertz, F Mormann, C Rieke, P David, CE Elger, , "Indications of nonlinear deterministic and finite-dimensional structures in time series of brain electrical activity" Dependence on recording region and brain state, Phys. Rev., E, vol.,64,pp.061907, 2001.



10.22214/IJRASET



45.98



IMPACT FACTOR:
7.129



IMPACT FACTOR:
7.429



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089  (24*7 Support on Whatsapp)