A Review on Epileptic Signals Analysis using KNN, K-means clustering and SVM Classifier

Sharad S. Jagtap
Assistant Professor
Department of Electronics & Telecommunication Engineering,
Anantrao Pawar College of Engineering & Research, Pune, India,

Abstract : for analysis of signals and its processing it is required to sort data for robust classification. Proposed paper represents classification of EEG signals using three different techniques. Brain waves of epilepsy are the data for this classification. Epilepsy is one type of neurological disorder which causes abnormal neuronal activity.
To identify the non linearity present in the data, Proper analysis is done. An EEG record of twenty patients is analyzed. The power spectral density is determined which is further used for disturbance reduction. The results show that SVM Classifier gives better results for classification of epileptic data as compare to KNN and K-means algorithm.

Index Terms **– KNN, SVM, Epilepsy data, EEG.**

I. INTRODUCTION

Human brain signal recording is important for both research purposes and assessment of various neurological disorders. For example, EEG is the current method of choice to visualize abnormal epileptiform discharges in patients with epilepsy. Continuous EEG monitoring is commonly used for the diagnosis and monitoring of convulsive or non-convulsive status epileptics and assessment of ongoing therapy for the treatment of seizures in such patients.

Some patients might get pleasure from brain disease surgery if the epileptogenic zone (EZ) will be known and reselected while not damage. There square measure 2 steps for epileptogenic zone localization: a) noninvasive, and b) invasive brain signal observance. The visual review of the brain signals provides a valuable diagnostic procedure for brain disease detection. Brain wave analysis will be performed by graph (EEG). alternative clinical tools on the market square measure generator encephalogram (MEG) and useful resonance Imaging (MRI).

Electroencephalogram is that the most utilized technique for seizure designation .During convulsion the onventional electrical activity of brain is discontinuous or altered which might be simply extracted with the assistance of electroencephalogram. this conductor arrangement follows commonplace 10-20 system, adopted by the Yankee electroencephalogram society. The signals no inheritable contain huge data which might be decoded to predict seizure.

The power spectral density is determined which is further used for dimensionality reduction. The performance index achieved by KNN classifier and K-means clustering are 78.31% and 93.02% respectively. A high Quality value of 22.37 with K-means clustering and a low value of 18.02 are obtained with KNN classifier. The results show that K-means outperforms KNN classifier in epilepsy risk level classification.

It uses a set of electrodes placed over scalp. This electrode arrangement follows standard 10-20 system, adopted by the American EEG society. The signals acquired contain enormous

Information which can be decoded to predict seizure [12]. In literature, there is a long list of methods that can be used to extract features from the EEG signals. These methods include Fourier Transforms (FT); good for analyzing stationary signals, and Time Frequency Distribution (TFD); provides a tool for examining continuous segments of EEG signals. However, EEG signals are non-stationary in nature and conventional frequency analysis methods may not capture the full details of the brain signals. Lyapunov exponents; discussed the detection and prediction of epileptic seizure in Refs. [5], analysis of correlation structure [6], and high order spectral analysis (HOS) [7] are the examples of non-linear methods for EEG signal analysis in the domain of epilepsy. Advances in wavelet theory has also made it a very suitable for bio-medical signal processing. It has a built-in advantage of capturing repeated and irregular patterns.

It may also trot out the analysis of transient and unforeseen signal changes [8,9]. this can be doable as a result of this system provides variable window size, slender at high and wide at low frequency levels. moreover [8], has mentioned many totally different strategies for EEG signal analysis and terminated that wave rework (WT) has additional blessings over different strategies.

II. METHODS

1. EEG Signal Capturing

A sixteen channel clinical encephalogram observance system, using 10-20 interactive electrode placement methodology is employed for getting the paper record of encephalogram. This raw encephalogram signals recorded square measure contaminated by artifacts. Major sources of artifacts square measure biological artifacts, contact electrical resistance from electrodes, and other electronic instruments within the area. With the help of specialist, the unit free encephalogram signals square measure selected and scanned by Umax 6696 Scanner with a resolution of 600 dpi. In this work, we've got used the encephalogram information set that has been no inheritable from twenty epileptic patients UN agency were underneath examination and drugs within the medical specialty Department of Sri Ramakrishna Hospital, Coimbatore. Using 5 pole Butterworth filter, the encephalogram signals square measure band passed between the frequency vary of zero.5Hz and 50Hz. The recorded encephalogram signals square measure continuous price of thirty second, they're reborn into epochs of 2 seconds. This two second epochs square measure enough for establish any medical specialty changes. the most frequency of encephalogram signals is fifty Hertz hence every epoch is sampled at two

hundred Hertz. every sample provides instantaneous amplitude of the encephalogram signals recorded. There are four hundred values for every epoch that contributes to massive dimension of information price. To eradicate this, spatiality reduction is finished.

2. Signal Conditioning

The preprocessing step is to scale back the dimension of the EEG data. the foremost typical spatial property reduction ways square measure Singular worth decomposition (SVD), Principal component analysis (PCA), and freelance part analysis (ICA) [2][13]. the downside of those techniques is that they can't handle advanced and non linear information [14]. In this work, we've used power spectral density (PSD) for the dimensionality reduction purpose. the ability spectral estimate of the whole EEG set is set. The power spectral density provides the data concerning the statistical distribution of EEG signals. It provides solely the magnitude data that is accessible within the frequency domain and doesn't offer any information concerning the part values.

3. Disturbance Analysis

Most of the biological signals measured are non linear in nature due to complexity of the system being measured [16]. To examine the non-linearity present in the epileptic EEG data deterred fluctuation analysis (DFA) is performed. The EEG signal recorded from scalp of the patient arises from non linear interactions of neurons. [17].The scaling exponent α is calculated for the EEG dataset. The autocorrelation properties of a signal are given below [18]

1. $\alpha < 0.5$ anti-correlated signal
2. $\alpha = 0.5$ uncorrelated signal (white noise)
3. $\alpha > 0.5$ positive correlation in the signal

4. Method 1-KNN Classifier

The KNN classifier may be a non constant quantity instance primarily based classifier [19]. it's a lazy learning methodology that doesn't learn from coaching information, merely stores all the samples within the training information. kept values are required throughout the training part. This algorithmic program relies on the closest neighborhood estimation. The new cases are classified on the basis of similarity live that is that the distance metric. Most commonly used is geometrician distance. disadvantage of KNN classifier is giant time needed to seek out nearest neighborhood in a giant coaching set [1]. thence spatiality reduction step is completed to beat this. In KNN classifier, the category of x is found by following procedure.

a) verify the k instances that are nearest to the
class x supported the gap live.
b) consecutive step is to permit this k instances to vote to
find the category of x.

5. Method 2- K-means clustering

K-Means clump is one in all the foremost widespread partitioned clustering algorithms [12]. it's a non settled and unsupervised partitioning methodology [15]. This algorithmic rule clusters the given set of knowledge into totally different subsets supported some criteria. the target of the algorithmic rule is to maximize the similarity between the patterns in same cluster whereas to minimize between totally different clusters [14].

The outline of basic K-means algorithm is given below.

a) Choose K points as the initial centriods.
b) Determine distance between each data point and the cluster centriod.
c) Reassign data point to new cluster with minimum distance.
d) Recomputed the new cluster centriods.
e) Repeat till convergence criteria is meet.

6. Method 3- SVM Classifier

The framework on that our system is predicated on is explicit as follows. a collection of better-known encephalogram recordings (for each seizures and traditional states) is given. for every recording, a window is softened over time in overlapping steps. For each frame of encephalogram signal, 3 options are extracted. When ail the frames inside a recording interval are processed, they're tagged consequently (-1 for epileptic seizures, +1 otherwise). once all the recordings within the set the set of options to fad the choice boundary that higher separates the 2 categories or classes. At now, the machine is claimed to be trained, i.e. it bas "learnt" the common options that characterize the epileptic seizures and is ready to spot them within the new recordings. The target values, i.e. -1 and +1, are given to the system throughout the coaching method. For the frequency analysis of the frames, the power spectral density should be calculable. this can be done by means that of associate degree autoregressive (AR) model to beat interference effect in DFT and therefore the restrictions iatrogenic by the nonstationarity of the signals. Moreover, AR filters use linear prediction of the output supported previous inputs, reinforcing the signal against the result of noise, which makes the ensuing spectra electric sander and fewer littered with noise. The order of the AR filter represents a compromise between accuracy and process speed. The prediction order will be calculable. For this application, values between ten to twelve are tried to be appropriate for OUT functions. The convenience for AR models and its concerns are mentioned in additional detail by different authors [10].

III. RESULTS

The relative performance of KNN classifier and K-means clustering are studied through calculating various parameters namely Performance index, Sensitivity, Specificity, Average detection and Quality value [26]. With K-means clustering, we have obtained the performance index of 93.02%, sensitivity of 100% , specificity of 93.75% and quality value of 22.37.The Quality value obtained with KNN classifier is less when compared to that of K-means clustering. The results obtained are tabulated below.

Performance Parameter	KNN method	K-Means Clustering	SVM Classifier
Accuracy (%)	83	93	99
Sensitivity (%)	90	100	100
Performance Index (%)	78	94	98
Specificity (%)	92	93	97
Time Delay (Seconds)	2.10	2.24	1.80
Quality Value(%)	18.02	22.37	42.50

IV. CONCLUSION

This paper investigates the performance of KNN classifier; K- means clustering and SVM in classifying the epilepsy risk level of epileptic patients from EEG signals. The parameters estimated are the power spectral density of the EEG signals and dimensionality reduction is carried out for these spectral values. Finally classification is done by applying KNN classifier, K-means clustering and SVM Classifier. Also the fluctuations present in the signals are studied. The aim was to design classification algorithm with high performance index and minimum false alarm rate and missed classification. We have obtained better performance with Support Vector Machine with false alarm rate of 0%. Thus applying this technique risk level of epilepsy can be identified and proper medication can be given to the patients.

REFERENCES

[1] G.O.Young,"Syntheticstructureofindustrial plastics,"in *Plastics*, 2nded., vol. 3, J. Peters, Ed. New York, NY, USA: McGraw-Hill,1964,pp.15–64.

[2] W.-K.Chen,*LinearNetworksandSystems*.Belmont, CA, USA:Wadsworth, 1993, pp. 123–135.

[3] J. U. Duncombe, "Infrared navigation—Part I: An assessment of feasibility," *IEEE Trans. Electron Devices*, vol. ED-11, no. 1, pp. 34–39, Jan. 1959,10.1109/TED.2016.2628402.

[4] E. P. Wigner, "Theory of traveling-wave optical laser," *Phys. Rev*., vol. 134, pp. A635–A646, Dec. 1965.

[5] E. H. Miller, "A note on reflector arrays," *IEEE Trans.Antennas Propagat.*, to be published.

[6] E. E. Reber, R. L. Michell, and C. J. Carter, "Oxygen absorption in the earth's atmosphere," Aerospace Corp., LosAngeles, CA, USA, Tech. Rep. TR-0200 (4230-46)-3, Nov. 1988.

[7] J. H. Davis and J. R. Cogdell, "Calibration program for the 16-foot antenna," Elect. Eng. Res. Lab., Univ. Texas, Austin, TX, USA, Tech. Memo. NGL-006-69-3, Nov. 15, 1987.

[8] *Transmission Systems for Communications*, 3rd ed., Western Electric Co., Winston-Salem, NC, USA, 1985, pp. 44–60.

[9] *Motorola Semiconductor Data Manual*, Motorola Semiconductor Products Inc., Phoenix, AZ, USA, 1989.

[10] G. O. Young, "Synthetic structure of industrial plastics," in Plastics, vol. 3, Polymers of Hexadromicon, J. Peters, Ed., 2nd ed. New York, NY, USA: McGraw-Hill, 1964, pp. 15-64. [Online]. Available: http://www.bookref.com.

[11] *The Founders' Constitution*, Philip B. Kurland and Ralph Lerner, eds., Chicago, IL, USA: Univ. Chicago Press, 1987. [Online]. Available: http://press-pubs.uchicago.edu/founders/

[12] The Terahertz Wave eBook. ZOmega Terahertz Corp., 2014. [Online]. Available: http://dl.z-thz.com/eBook/zomega_ebook_pdf_1206_sr.pdf. Accessed on: May 19, 2014.

[13] Philip B. Kurland and Ralph Lerner, eds., *The Founders' Constitution.* Chicago, IL, USA: Univ. of Chicago Press, 1987, Accessed on: Feb. 28, 2010, [Online] Available: http://press-pubs.uchicago.edu/founders/

[14] J. S. Turner, "New directions in communications," *IEEE J. Sel. Areas Commun.*, vol. 13, no. 1, pp. 11-23, Jan. 1995.

[15] W. P. Risk, G. S. Kino, and H. J. Shaw, "Fiber-optic frequency shifter using a surface acoustic wave incident at an oblique angle," *Opt. Lett.*, vol. 11, no. 2, pp. 115–117, Feb. 1986.

[16] P. Kopyt*et al.,* "Electric properties of graphene-based conductive layers from DC up to terahertz range," *IEEE THz Sci. Technol.*, to be published. DOI: 10.1109/TTHZ.2016.2544142.

[17] PROCESS Corporation, Boston, MA, USA. Intranets: Internet technologies deployedbehindthefirewall forcorporateproductivity. PresentedatINET96AnnualMeeting.[Online].Available:http://home.process.com/Intranets/wp2.htp

[18] R. J. Hijmans and J. van Etten, "Raster: Geographic analysis and modeling with raster data," R Package Version 2.0-12, Jan. 12, 2012. [Online]. Available: http://CRAN.R-project.org/package=raster

[19] Teralyzer. Lytera UG, Kirchhain, Germany [Online]. Available: http://www.lytera.de/Terahertz_THz_Spectroscopy.php?id=home, Accessed on: Jun. 5, 2014