ISSN:0975-9646

Priyanka Jaiswal et al, / (IJCSIT) International Journal of Computer Science and Information Technologies, Vol. 6 (3) , 2015, 2031-2034

EEG signal classification using Modified Fuzzy Clustering algorithm

Priyanka Jaiswal, Rupali Koushal

Information Technology, Shri Vaishnav Institute of Technology and Science
Ujjain road, Gram Baroli, Indore, M.P.,India

***Abstract*— Epilepsy is a brain disorder in which clusters of nerve cells, or neurons, in the brain sometimes signal abnormally. The Empirical Mode Decomposition (EMD) is used to extract the features of EEG signals which help us to detect the epilepsy. In this paper, An Enhanced Classifier with Modified Fuzzy Clustering Algorithm to detect epilepsy is proposed. This proposed approach is evolving for multiclass classification problem. Bayesian theory is utilized to formulate the problem of clustering and classification. In clustering algorithm the selection of learning parameter i.e., clusters membership Degree is initially chosen at random, but here in the proposed methodology, the value of clusters membership degree is calculated on the basis of randomly initialized cluster centers. It is demonstrated by experiments that, this method improve the performance of the algorithm. The same is being verified with EEG signal datasets.**

Keywords—Epilepsy, EEG signal, Emperical Mode Deccomposition.

I. INTRODUCTION

This document is a template. An electronic copy can be downloaded from the conference website. For questions on paper guidelines, please contact the conference publications c Pattern recognition is the assignment of some sort of output label to a given input label, by applying some machine learning algorithm to it. Watanabe [2] defines an entity as a pattern that could be given a name. Pattern recognition techniques are widely used in various application domains such as hand writing recognition and data mining [3] [4]. Given a sample, its recognition may consist of one of the following two tasks: unsupervised clustering and supervised classification. Supervised classification assigns new sample to a class based on the feature of this sample i.e., classifiers designed in supervised manner directly employ the class labels to learn the classification decision function which in turn used to label a new sample. Because of this, these classifiers rarely cares about the revelation of the data structure and just emphasize the determination of the decision functions, but As a result, these classifiers fail to understand the classification result.

Electroencephalography (EEG) [1] is the recording of electrical activity which contains information about human brain functionality and the disorders of the nervous system. EEG, or electroencephalograph, deserves mention as one of the rest ways of non-invasive observing activity of the brain activity. An EEG is a recording of electrical signals from the brain. It is done by fixing up electrodes to the patient’s scalp. EEG accurately analysis the deviations of electric signals through small interval through different electrodes placed on the patient scalp, the changes in these electric signals are measured in terms of voltage fluctuations of brain. The information about the human brain and neurological disorders is found through the output of the electrodes. An EEG can show the state of a person such as numb, awake, asleep because the characteristic patterns of current vary for the aforementioned states.

Epilepsy [3] is a brain disorder in which clusters of nerve cells, or neurons, in the brain sometimes signal abnormally. In epilepsy, the normal pattern of neuronal activity becomes disturbed, causing strange sensations, emotions, and behavior, or sometimes convulsions, muscle spasms, and loss of consciousness. The Epilepsy is known by disaster function of the brain which is termed seizure. Ictal and Interictal [2] are the medical conditions of seizure, where the period of the seizure is represented by Ictal and the intermediate period between two seizures is represented by Intericatal. However we have to make a note that Interictal di ers from that of a non-seizure signal. A prediction of the Ictal from Interictal could make the patient to put away from the next seizure. It is imprecise and erroneous to detect epilepsy by visual scanning of EEG Signals. The detection of epileptic seizures, which are convulsions accompanied by impaired consciousness, in the EEG signal is a vital part in the diagnosis of epilepsy. Nonetheless, Classification between the ictal and interictal is essential for the detection of Epileptic seizures.

Classification of EEG signal is very time consuming task because a human expert is required to classify them properly. So, in this paper, we present a Novel An Enhanced Classifier with Modified Fuzzy Clustering Algorithm to detect epilepsy is proposed. In this paper, first we extract the features of EEG signal which help us to classify the epilepsy problem. The rest of the paper is organized as follows: Section II contains the literature review. Section III contains the proposed work. Section IV contains the experimental results and Section V concludes our work.

II. LITERATURE REVIEW

M. Setnes and R. Babuska [6] proposed Fuzzy relational classifier (FRC) in which they used a fuzzy logic relation to establish the relation between structures in the feature space and their corresponding class labels. Fuzzy relational classifier is especially useful for the problems where one does not have a prior knowledge that which features are selected to describe about the class. Fuzzy logic relation solves the problem of labeling the prototypes, which can be particularly difficult when the training data contains classification errors. The fuzzy relational based classification scheme provides transparency for a complex nonlinear classification problem. The transparency of relational classifier provides better analysis of both the training as well as classification results for unseen patterns. The training phase of the classifier proceeds in two phases. Firstly, in order to explore the natural structure of samples the clustering is performed in unsupervised manner and then the resulted cluster membership is used to establish the fuzzy relation matrix R corresponding to each sample.

M. Setnes and R. Babuska [6] proposed Fuzzy relational classifier (FRC) in which they used a fuzzy logic relation to establish the relation between structures in the feature space and their corresponding class labels. Fuzzy relational classifier is especially useful for the problems where one does not have a prior knowledge that which features are selected to describe about the class. Fuzzy logic relation solves the problem of labeling the prototypes, which can be particularly difficult when the training data contains classification errors. The fuzzy relational based classification scheme provides transparency for a complex nonlinear classification problem. The transparency of relational classifier provides better analysis of both the training as well as classification results for unseen patterns. The training phase of the classifier proceeds in two phases.

Firstly, in order to explore the natural structure of samples the clustering is performed in unsupervised manner and then the resulted cluster membership is used to establish the fuzzy relation matrix R corresponding to each sample. W. L. Cai, S. C. Chen, and D. Q. Zhang proposed Enhanced Fuzzy Relational Classifier (EFRC) they build more robust relation matrix R, to overcome the drawback of class-overlapping regions that leads to the formation of impure cluster. The problem of class-overlapping regions is eliminated indirectly by constructing the representative subset of samples by purifying the formed clusters from the original training dataset. Hence the representative subset so obtained with the purer cluster enhances the consistent relation between the formed clusters and their dominant classes correspondingly.

The relation matrix R so constructed using the EFRC classification scheme reflect the better relationship between the clusters and its corresponding classes and leads to enhance the robustness and reliability of the achieved classification performance. EFRC classifier suffers from a major drawback because of the use of a non- robust Euclidean distance that makes the classification scheme sensitive towards the noise and outliers.

W. L. Cai, S. C. Chen, and D. Q. Zhang proposed Robust Fuzzy Relational Classifier (RFRC) to improve the robustness against outliers and by providing suitability for the non-spherical data structure. In RFRC algorithm, firstly the kernelized fuzzy c-means algorithm (KFCM) is adopted in order to replace the non-robust Euclidean metric used in the objective function of FCM. It is due to the robust distance metric induced by the kernel function. It is because of robust distance metric in the KFCM algorithm it is proven to be robust and also able to group the non-spherical clusters from the given dataset. Then, in the second training step of RFRC, the fuzzy set theory is applied to compute the soft class labels which will more precisely described the class information associated with the data sets.

Consequently, by incorporating both the KFCM and the soft class labels the relation matrix R is constructed which reflect the fuzzy relation between clusters and classes for the subsequent classification, thus by significantly decreasing the reject rate and by boosting the accuracy. The RFRC classifier suffers from a disadvantage because the relational matrix R is constructed using composite operator as follows: 1) In RFRC classifier, the optimization of the fuzzy relational matrix R is difficult due to the in differentiability and complexity of the composite operators 2) the entries in relation matrix R also lacks a statistical characterization and thus fail to indicate the reliability of relationship between the formed cluster and its corresponding classes 3) Due to the presence of inconsistent class information in the training dataset, the entries in relation matrix R approaches to 0 and thus fails to reflect any discriminate information.

Tomasz M. Rutkowskiy et al. proposed a application of a machine learning algorithm based on the multivariate empirical mode decomposition (MEMD) method to EEG signal separation. The automatic multivariate intrinsic mode functions (IMF) clustering allowed us to distinguish the seizure related spikes from the normal ones. This method has been compared with classical blind separation approach based on ICA, which failed to identify the EEG signals related to the brain seizures.

Recently, nonlinear methods have been proposed to extract parameters for analysis and classification of EEG signals. Among these parameters, the Lyapunov exponent provides clinically useful information, and it can be useful in analysing long term EEG signals for early detection of electroencephalographic changes. The correlation dimension allows a quantitative description of complexity changes of neuronal behaviour of the human brain . The fractal dimension parameter measures the complexity, irregularity, and the chaotic nature of the EEG signals, and has provided discrimination between normal and epileptic seizure EEG signals . The approximation entropy (ApEn) is a recently formulated statistical parameter to quantify the regularity of EEG signals. It has been demonstrated that value of the ApEn drops abruptly due to synchronous discharge of large groups of neurons during an epileptic activity.

A complexity analysis and spectrum analysis features have been utilized for the classification of EEG signals for epilepsy detection . Newer techniques for analysis of EEG signal for the detection of epilepsy have been proposed in , which are based on the empirical mode decomposition (EMD) developed especially for nonlinear and non stationary signal analysis.

III. Proposed Work

A. Data Processing

This section shows how we extract features from EEG signals to use as an input for GP. Empirical Mode Decomposition The empirical mode decomposition (EMD) method was developed by Huang et al. [8] to decompose functions into a superposition of natural modes, each of which could be easily analysed for their instantaneous frequencies and bandwidths. EMD is basically a method of breaking down a signal without leaving the time domain introduced for analysis of nonlinear and non- stationary signals. It can be compared to other analysis methods like Fourier Transforms and wavelet decomposition. The algorithm includes the following steps:

1) Calculate the IMF for each iteration using EMD on EEG signals.

2) Calculate two features namely Frequency parameter and Amplitude parameter using Hilbert transform applied on IMF's for each iteration.

3) Generation a Bandwidth parameter by combining frequency parameter and Amplitude parameter.

B. Calculation of Intrinsic Mode Function

Using the EMD algorithm, we obtain intrinsic mode functions (IMF), which were generated at each scale, going from one to coarse, by an iterative procedure to locally isolate the modal behaviour. In contrast to the aforementioned Fourier transform and wavelet transform, the EMD decomposes any given data into intrinsic mode functions (IMF) and a residual function that are not set analytically and are instead determined by an analysed sequence alone. The basic functions are in this case derived adaptively directly from input data. The IMF resulting from the EMD shall satisfy only the following requirements:

1) The number of IMF extreme (the sum of the maxima and minima) and the number of zero-crossings must either be equal or differ at most by one

2) At any point of an IMF the mean value of the envelope defined by the local maxima and the envelope defined by the local minima shall be zero. The algorithm as proposed by Huang requires the identification of all local extrema that are further connected by cubic spline lines to produce the upper and the lower envelopes

C. Proposed Algorithm

We will classify the EEG signal with the Modified Fuzzy clustering Algorithm. In this algorithm:

1) Initially we have training dataset, class membership of training dataset.

2) Than compute fuzzy cluster centres randomly.

3) Apply the formula of Euclidian Distance

$$D_{ij}=\sqrt{\sum(x_i-c_j)^2}$$

3.1) Creating membership metrics by using formula

$$U_j(x_i)= (1/d_{ji})^{1/m-1} / {}_{k=1}\sum^{p} (1/d_{ki})^{1/m-1}$$

3.2) Generating the new centred

$$C_{j=}\ {}_{i=1}\sum[u_j(x_i)]^2 x_i \ / \ {}_{i=1}\sum[u_j(x_i)]^2$$

4) Repeat step 3 till optimal cluster centre is achieved.

5) Calculate P relationship matrix using cluster membership and given class membership.

IV. EXPERIMENTAL RESULT

A. *Dataset*

An EEG dataset, which is publicly available online in [27] is used in this work. Comparison of Classification accuracies and training time.

In this section, the classification performance comparison is made among FRC, RFRC, SCC and proposed method. In all of our experiments, each data set is randomly partition into two halves: one half is used for training and the other half for testing and results is generated in terms of accuracy as shown in Table I.

TABLE I. COMPARIOSN OF CLASSIFICATION ACCURACY ON EEG DATASET

Dataset	**Methods**			
	FRC	***RFRC***	***SCC***	***Proposed***
EEG	69%	72%	75%	90%

We also compare the training time of all these methods and the result are presented in Table II.

TABLE II. COMPARIOSN OF TRAINING TIME ON EEG DATASET

Dataset	**Methods**			
	FRC	***RFRC***	***SCC***	***Proposed***
EEG	69%	72%	75%	90%

It is clear from Table I that our proposed method is able to achieve better classification accuracy among all the above methods. It is mainly due to the updating we had done in our methods which help us to reach the solution fast. It is also clear from Table II that our method take very less time in comparison to other methods for classifying the EEG dataset which again shows the superiority of our algorithm over other algorithms.

V. CONCLUSIONS

In this paper, An Enhanced Classifier with Modified Fuzzy Clustering Algorithm to detect epilepsy is proposed. It consists of two steps. In the first step, clustering mechanism is designed to make the clusters of data.

In the second step, a proper relation is built that maps the relationship between the cluster membership and the class membership. It aims to simultaneously optimize the clustering and classification criteria without sacrificing the clustering and classification performance.

Clustering and classification results can also be yielded by optimizing both the clustering centres and the covariance used in the Multiobjective functions. The proposed approach has been applied on EEG dataset, illustrates that the classification accuracy achieved by our method is higher as compared to the FRC, RFRC, SCC due to the proposed framework which significantly improve the clustering criteria by effectively minimizing the number of iteration required to calculate the cluster center along with the better interpretation of classification results.

ACKNOWLEDGMENT

The authors express their thanks to Assistant Prof Priyanka Gupta, SVITS, Indore, who has helped in different ways to complete this work. Thanks are also due to the anonymous reviewers and our colleagues for their suggestions to improve the manuscript.

REFERENCES

[1] Weiling Cai, Songcan Chen, and Daoqiang Zhang, "A Multiobjective Simultaneous Learning Framework for Clustering and Classification" IEEE transactions on neural networks, vol. 21, no. 2, February 2010.
[2] S.watanabe, Pattern Recognition: Human and Mechanical. New York: Wiley, 1985.
[3] A.K. Jain, R. P. W. Duin, and J. Mao, "Statistical pattern recognition: A review," IEEE Trans. Pattern Anal. Mach. Intell., vol. 22, no. 1, pp.4–37, Jan. 2000.
[4] M. Setnes and R. Babu..ska, "Fuzzy relational classifier trained by fuzzy clustering," IEEE Trans. Syst. Man Cybern. B, Cybern., vol. 29, no. 5, pp. 619–625, Oct. 1999.
[5] W. L. Cai, S. C. Chen, and D. Q. Zhang, "Enhanced fuzzy relational classifier with representative training samples," presented at the Int.Conf. Wavelet Anal. Pattern Recognit. Beijing, China, Nov. 2–4, 2007
[6] W. L. Cai, S. C. Chen, and D. Q. Zhang, "Robust fuzzy relational classifier incorporating the soft class labels," Pattern Recognit. Lett.,vol.28, pp. 2250–2263, 2007.
[7] W. L. Cai, S. C. Chen, and D. Q. Zhang, "A simultaneous learning framework for clustering and classification,"PatternRecognit.2009 [Online]. Available:http://www.dx.doi.org/10.1016/zj.patcog, acceptedfor publication.
[8] A. C. Coello, G. T. Pulido, and M. S. Lechuga, "Handling multiple objectives with particle swarm optimization," IEEE Trans. Evol. Comput., vol. 8, no. 3, pp. 256–279, Jun. 2004.
[9] J. Clerk Maxwell, A Treatise on Electricity and Magnetism, 3rd ed., vol. 2. Oxford: Clarendon, 1892, pp.68-73.
[10] I.S. Jacobs and C.P. Bean, "Fine particles, thin films and exchange anisotropy," in Magnetism, vol. III, G.T. Rado and H. Suhl, Eds. New York: Academic, 1963, pp. 271-350.
[11] Y. Yorozu, M. Hirano, K. Oka, and Y. Tagawa, "Electron spectroscopy studies on magneto-optical media and plastic substrate interface," IEEE Transl. J. Magn. Japan, vol. 2, pp. 740-741, August 1987 [Digests 9th Annual Conf. Magnetics Japan, p. 301, 1982].
[12] M. Young, The Technical Writer's Handbook. Mill Valley, CA: University Science, 1989.