IJCSE International Journal of Computer Sciences and Engineering Open Access
Research Paper Vol.-7, Issue-1, Jan 2019 E-ISSN: 2347-2693

K-modes and Fuzzy C-means with modified Particle Swam Optimization Clustering Algorithm for Epilepsy Seizure Data

C.V. Banupriya[1*], D. Deviaruna[2]

[1]Dr.NGP Arts and Science College, Department of Computer Science, Coimbatore, India
[2] Dr.NGP Arts and Science College, Department of Computer Applications, Coimbatore, India

[*]*Corresponding Author: banupriya.venkat@gmail.com*

Available online at: www.ijcseonline.org

Accepted: 19/Jan/2019, Published: 31/Jan/2019

***Abstract*—** Epilepsy is a stable neurological disorder of the brain, described by regular seizures, i.e., irregular activities. Seizure is the most imperative signal of epilepsy, which is solitary of the most expected neurological disorders. An electroencephalogram (EEG) is a test out used to weigh up the electrical activity in the brain, and is widely used in the recognition and study of epileptic seizures. Hence, it is decisive to develop a quantitative technique to automatically clustering the normal and epileptic brain activities. Several techniques have been developed for unbending out the important features of seizures present in EEGs. The proposed approach is evaluated an extracting the features of EEG signals using wavelet transform coefficients and unsupervised learning technique like clustering the data using Fuzzy C- Means with Modified Particle Swarm Optimization (PSO) and K- Mode Clustering. The recital of the Clusters are analyzed and examined that Fuzzy C-Means with PSO less error rate and out performs than K-Mode Clustering in accuracy.

***Keywords*—** Fuzzy C-means, K-mode, EEG, Seizures, Wavelet, PSO, Clustering

I. INTRODUCTION

Epilepsy is a stable neurological disorder of the brain that affects more or less 50 million people worldwide. According to the World Health Organization (WHO), epilepsy is characterized by frequent seizures, which are goal reactions to sudden, usually succinct, too much electrical discharges in a group of brain cells [1]. In the structure of epilepsy examining, two various categories of seizure cover to be believed, namely behavioral and electrographic. A behavioral seizure is characterized as the medical manifestations of epilepsy, as assumed by the patient, seen by an observer, or recorded on video. An electrographic (or electroencephalographic (EEG)) seizure is identified as an abnormal bursting EEG pattern. In numerous cases, there is dissociation between behavior and EEG signals [2]. In sustain types of epileptic seizures can occurred at all ages. Common epileptic seizures can be subdivided into nonexistence (petit mal) and tonic-clonic seizures [3].

In epilepsy the regular pattern of neuronal movement becomes disturbed, causing weird feelings, feelings, and actions, or occasionally convulsions, muscle spasms and loss of perception [4]. There are mixtures of achievable sources of epilepsy. Something that alerts the normal sample of neuron activity ranging from illness to brain damage to abnormal brain improvement can guide to seizures [5].

Standard EEG seizure finding, quantification, and detection have been areas of agonize and learning within the medical, physics, and engineering communities since the decades. In the most recent pair of years, the EEG scrutiny was mostly alert on epilepsy seizure recognition diagnosis. A seizure recognition configuration can be alienated into three stages: data acquirement and preprocessing, processing and feature extraction, and clustering.

The rest of the paper is organized as pursues: related work is discussed in Section II. Methodology discussed in Section III. The experimental results are shown in Section IV. The Section V Concludes the paper based on experimental Clustering.

II. RELATED WORK

Aarabi, R et.al [6], in biomedical signal processing, it is serious to determine the echo and pieces current in the raw signals so that their power in the feature extraction part can be minimized. Intracranial EEG (iEEG) signals were band-pass-filtered between 0.5 and 100 Hz to only allow the frequencies of concentration and a mark filter was used to eradicate 50-Hz power stroke sound. Transmission and stress group objects were identified. Divisions for which the derivative of the iEEG signal was zero were outstanding as having infiltration objects. Stress group objects were

removed using an entrance value; Intracranial EEG separations containing a signal with amplitude of healthier than 1.5mV was considered as having stress group objects.

H. Adeli et.al [7], a wavelet filter that requires the frequency at easiness to be limited to the 0-60 Hz group is used, so the EEG is group limited to the favored group by convolving with a low-pass finite impulse response (FIR) filter.

S. Tong et.al [8], Independent component analysis (ICA) is predominantly used for object cancellation. ICA identifies causes; in this folder an object near in the EEG signal, based on screen foundation separation (BSS) and separates those from the EEG based on their geometric independence. Varsavsky. I et.al [9], the pre-processing portion also normalizes the signal to formulate the data correspondent with those recorded by another gaining structure or from a dissimilar tolerant. An illustration of this is locating all the data in given amplitude series, permitting the signals to be computes up to directly.

Swami et al. (Swami et al., 2016) take out hand-crafted attributes such as Shannon entropy, standard deviation, and energy. They use up the common regression neural network (GRNN) classifier to classify these aspects and achieved maximum precision for A-E (normal vs. seizure) and AB-E (normal vs. seizure) folders, correspondingly on University of Bonn dataset. However, maximum accuracy for other folders likes B-E, C-E, D-E, CD-E, and ABCD-E. In additional research learning, Guo et al. (Guo et al., 2010) attained the maximum accuracy for ABCD-E folder on the similar dataset. They applied artificial neural network classifier (ANN) to classify the line length elements that were take out by using discrete wavelet transform (DWT).

Nicolaou et al. (Nicolaou et al., 2012) extracted the conversion entropy attribute from EEG signals. They applied support vector machine (SVM) as a classifier and attained the maximum accuracy for A-E folder on the University of Bonn dataset. However, highest accuracy for other folders such as B-E, C-E, D-E, and ABCD-E .Gandhi et al. (Gandhi et al., 2011) extorted the entropy, standard deviation and energy attributes from EEG signals using DWT. They employed SVM and probabilistic neural network (PNN) as a classifier and explained the maximum accuracy for ABCD-E folder.

III. METHODOLOGY

A. K- Mode:

The K-means cluster process cannot cluster unqualified data since of the disparate assess it using. The K-mode clustering algorithm is place on K-mean prototype other than get rid of the numeric data constraint even as shield its efficiency. This K-mode technique broaden K-mean mold to cluster unqualified data.

The K-modes [10] recognize that the in sequence of amount of feasible group of data (i.e. K) is available and consists of the subsequent paces: -

1. Create K clusters by subjectively selecting data substances and prefers K preliminary cluster core, one for each of the cluster.

2. Dispense data entity to the cluster whose cluster core is close to in the direction of it.

3. Update the K cluster support on distribution of data objects boon Calculate K newest modes of each one clusters.

4. Reiterate step 2 to 3 in anticipation of no data entity has altered cluster rapport or else various supplementary predefined measures are fulfilled [11].

B. Fuzzy C- Means with Modified Particle Swarm Optimization (PSO):

The crucial FCM algorithm is earliest enlightened for a set of data positions and then the algorithm is enlarged to a gray-scale picture [12].

Let us imagine we have a set of data positions a = {a1, a2, a3 … ai} where both point aw = {a1, a2, a3… aj} is a j dimensional vector. Precisely, they form a j dimensional vector and so on i * j template.

Let us visualize that environment is A[i][j]; Let us articulate, we have 'K' clusters. So the trouble becomes dividing the statistics into k clusters such that the space among the centroid and apiece position is minimum.

Here once more, the space may be Euclidean space (difference among the square of spaces and the particular position) or the Manhattan space (modulus of space among the position and the centroid).

Before we go profound into FCM, we want to be glad about the idea of fuzzy relationship and its position in clustering. Each position in a cluster is allocated a relationship supported on the possibility of that position be in the right place to the cluster.

For instance a cost "23" is rationally near both 20 and 25. I utilize the word "rationally" at this time in the fuzzy perception. If we glance at the distance between the location 23 to 20 and 25, 23 is nearer to 25. But a situation force happens when 23 needs to fit in to the clusters where 20 and 25 are centroid.

Therefore, the possibility of 23 belonging in 25 is declare, 0.8 then the possibility of 23 belonging to 20 is 0.2 haughty there are presently two clusters.

The condition of leased space is only one such instance for the fuzzy C - means clustering [13].
At present the FCM algorithm turn into an optimization dilemma with minimizing space among the position and the cluster centre known as the "limitation". Having seized the essentials to a rational amount, let us now go to the arithmetic of the entire problem.

The vital algorithm paces for FCM are,
1. Allocate a primary random centroid to every cluster (Faction).

2. Calculate the distance among both position and the cluster centre using an uncomplicated algorithm.

3. Supported on space among both position and the cluster centre, re-calculate the relationship function.

4. Supported on the new relationship function, re-calculate the centroid.

5. If the difference among the original centroid and the subsequently one is lower than an assured threshold value then the algorithm prevents, in addition it persists plow this stipulation is true.

We can acquire improved initial clustering centroid using several other methods and thus it may express Fuzzy C-means algorithm to accomplish superior in the progression of finding the superlative probable clustering centers by the mode of humanizing the clustering centroid.

The Modified Particle Swarm Optimization (PSO) algorithm is a heuristic method that optimizes a predicament by iteratively demanding to progress a contestant answer with observe to a given calculate of excellence. The enhanced initial cluster centroid can be attained using Modified PSO.

We propose Modified PSO combined with Fuzzy C-means algorithm that fabricates a speedy and improved clustering and also it tries to keep away from in getting mesmerized towards a restrained best potential outcome.

IV. EXPERIMENTAL RESULTS

To estimate the efficiency of k-mode and fuzzy C-means with modified PSO performance carry out research on EEG Seizure Detection dataset.

Object in an ith cluster are assumed to be classified every correctly or incorrectly through admiration to a famous class of object.

The clustering accuracy, sensitivity, specificity, error rate, precision and fscore are shown in Table 1 as follows.

The proportional analysis of the parameter evaluation in MATLAB 2013 is exposed in the subsequent figures.

Table 1: Parameter Measures

Parameter Measures	K Mode	Fuzzy C – Means with Modified PSO
Accuracy	0.49	0.92
Sensitivity	0.20	0.92
Specificity	0.81	0.91
Precision	0.53	0.92
Error Rate	0.50	0.07
Fscore	0.29	0.92

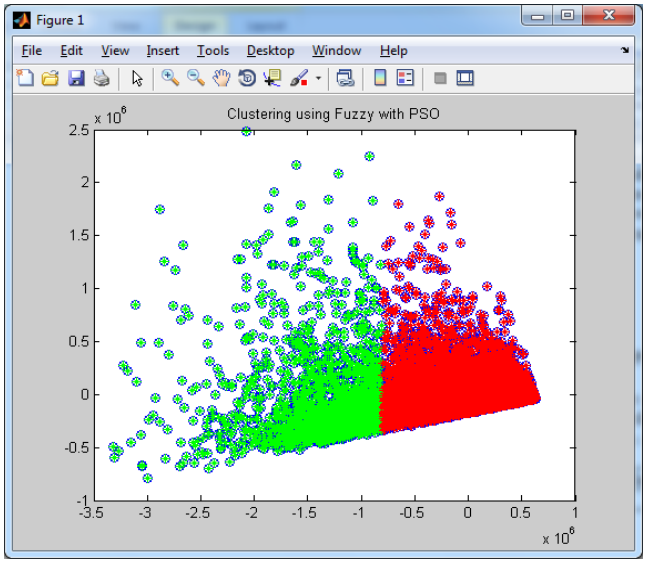

Figure 1. Results of Clustering Using Fuzzy with PSO

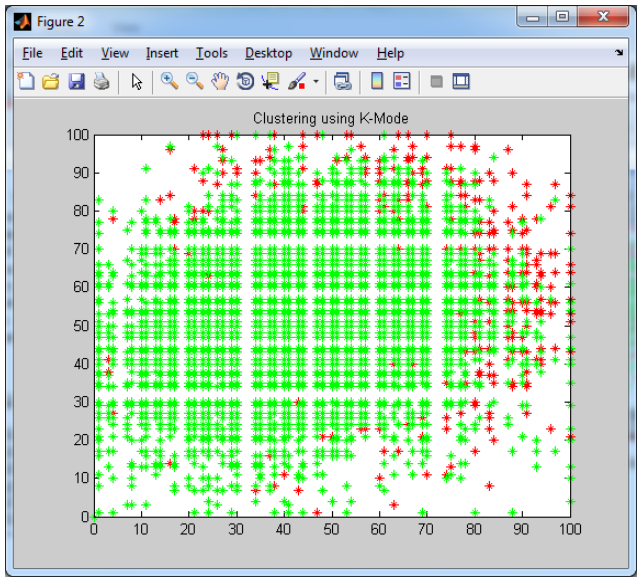

Figure 2. Results of Clustering Using K-Mode

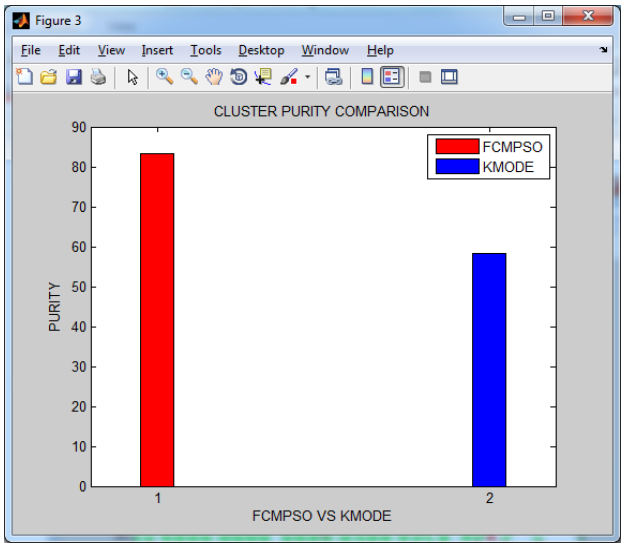

Figure 3. Results of Cluster Purity Comparison

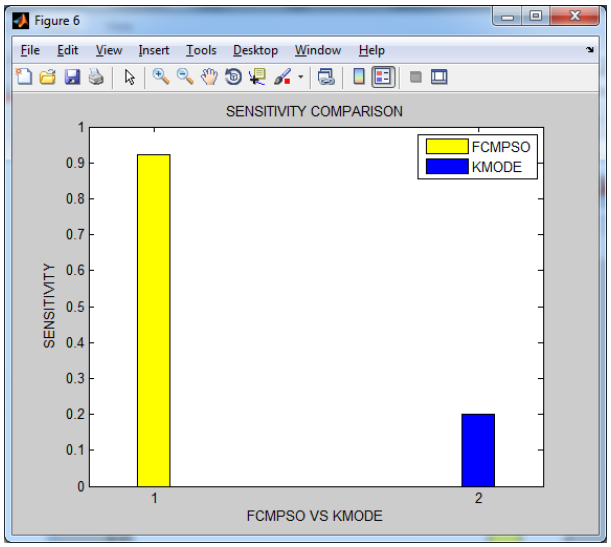

Figure 6. Results of Cluster Sensitivity Comparison

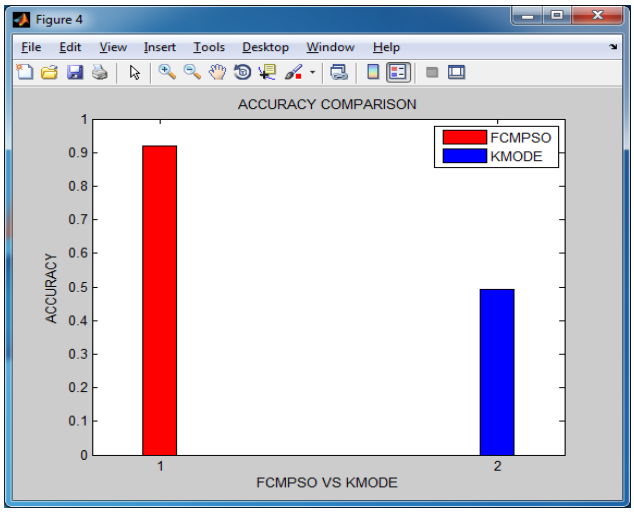

Figure 4. Results of Cluster Accuracy Comparison

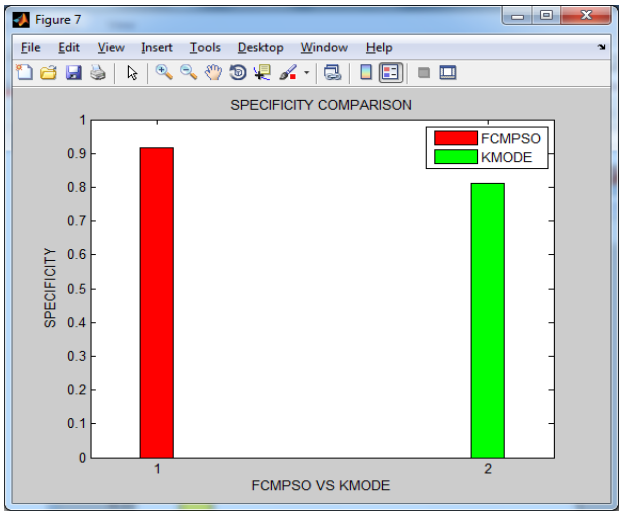

Figure 7. Results of Cluster Specificity Comparison

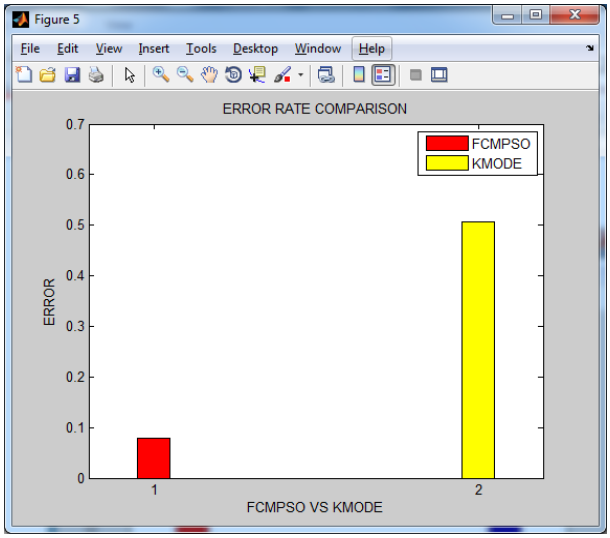

Figure 5. Results of Cluster Error Rate Comparison

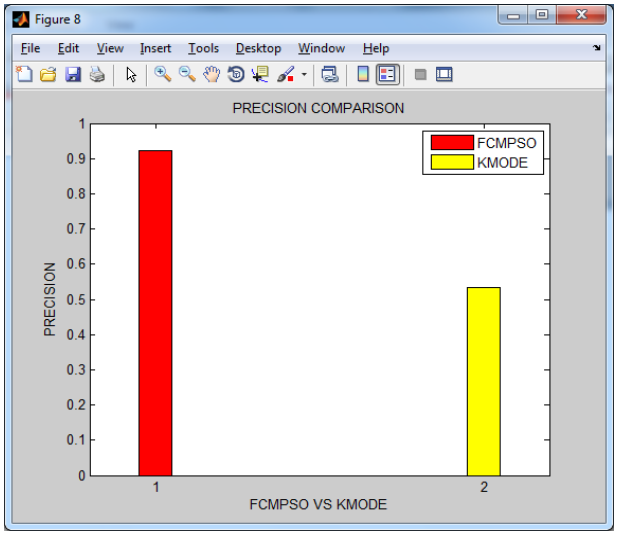

Figure 8. Results of Cluster Precision Comparison

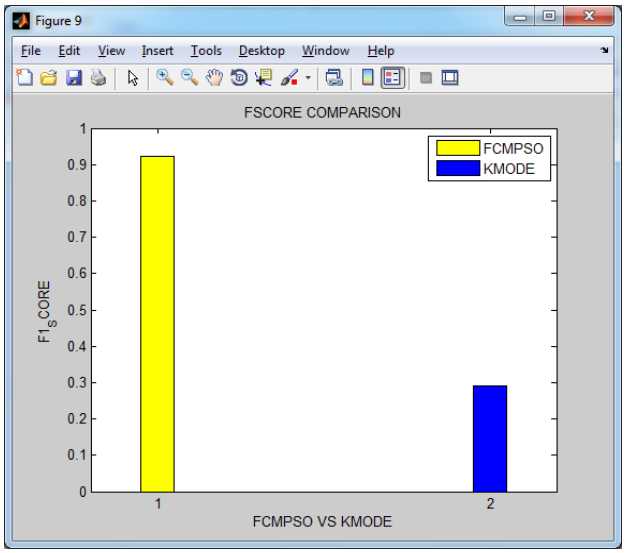

Figure 9. Results of Cluster Fscore Comparison

V. CONCLUSION

Epilepsy is a consistent neurological disorder of the brain that affects more or less peoples in this world. Day by day brain disorder patients are rising in all over the world. The proposed work comparing the Epilepsy Seizure classification task using powerful unsupervised learning algorithms like clustering techniques namely K-Mode and Fuzzy C-means with modified PSO. The outcome of the experiments indicates that the Fuzzy C-means with modified PSO clustering is more accurate and faster than the K-Mode Clustering Technique. The primary target of this work is to Accumulate Purity of the Seizure Detection Dataset.

REFERENCES

[1] World Health Organization, 2011.Available: http://www.who.int/mediacentre/?factsheets/fs999/en/.
[2] J. Gotman, "Automatic detection of seizures and spikes," J. Clin. Neurophysiol. 16: 130-140, 1999.
[3] James, C. J. "Detection of epileptic form activity in the electroencephalogram using the electroencephalogram using artificial neural networks", University of Canterbury, Christchurch, 1997.
[4] J.D. Bronzing, "Biomedical Engineering Handbook", New York: CRC Press LLC, Vol. I, 2nd edition 2000.
[5] Hojjat Adeli, Samanwoy Ghosh and Dastidat, "Automated EEG Based Diagnosis of Neurological Disorders", CRC Press; 1 edition, 2010.
[6] Aarabi, R. Fazel-Rezai and Y. Aghakhani, "A fuzzy rule-based system for epileptic seizure detection in intracranial EEG," Clin. Neurophysiology, 120: 1648-1657, 2009.
[7] H. Adeli, S. Ghosh-Dastidar and N. Dadmehr, "A Wavelet-Chaos methodology for analysis of EEGs and EEG sub bands to detect seizure and epilepsy," IEEE Trans. Biomed. Eng., 54: 205-211, 2007.
[8] S. Tong and N. V. Thakor (Ed.), "Quantitative EEG Analysis Methods and Clinical Applications", Norwood: Artech House, 2009.
[9] Varsavsky, I., Mareels and M. Cook, "Epileptic seizures and the EEG", Boca Raton: CRC Press, 2011.
[10] Zhexue Huang, A Fast Clustering Algorithm to Cluster Very Large Categorical Data Sets in Data Mining.
[11] Guo Tao, Ding Xingu, Li Yefeng, Parallel *k*-modes Algorithm based on MapReduce.
[12] S. Agrawal, R. Panda, L. Dora, "A study on fuzzy clustering for magnetic resonance brain image segmentation using soft computing approaches", Applied Soft Computing, 2014.
[13] D. L. Pham, "Fuzzy Clustering With Spatial Constraints", Image Processing. 2002. Proceedings. 2002 International Conference, 2002.

Authors Profile

Mrs. C V Banupriya Bachelor of Science from Madurai Kammaraj University in 2003, Master of Computer Application from Madurai Kammaraj University in year 2006 and Master of Philosphy in Computer Science from Bharathiar University in year 2014. She is currently pursuing Ph.D. and currently working as Assistant Professor in Department of Computer Science, Bharathiar University affiliated College since 2014. She has published many research papers in reputed international journals and conferences including Springers and it's also available online. Her main research work focuses on Data Mining. He has 4 years of teaching experience and 3 years of Industrial Experience.

Mrs Dr.Deviaruna.D currently working as Assistant Professor in Department of Computer Applications, in Dr.NGP College of arts and science. She has published many research papers in reputed international journals and conferences including IEEE and it's also available online. Her Specialization is Networks.

