

## EVALUATION AND DIAGNOSIS OF ADHD WITH K-MEANS AND FUZZY K-MEANS DATAMINING CLASSIFIERS

*Neema.H.N<sup>1</sup>, Dr. N.V.Balaji<sup>2</sup>*

### ABSTRACT

Attention Deficit Hyperactivity Disorder (ADHD) is a disorder among children which needs an early diagnosis. The projected work comprises disorder factors, demonstration, evaluation and diagnosis process. This issue is prevalent among many children and a great burden on their parents. Hence, the method has drawn interest from the areas of both health and education. The ADHD is a cerebral disorder which has more impact on school children's life and this becomes evident they have difficulty in controlling their behavior and concentrating on anything for long. Hence these types of children can be classified based on two types such as: a) Non – ADHD and b) ADHD. The data mining algorithms will do a significant job in classifying the types. The existing approach deals with the K-means clustering to perform classification. But the existing method faces more problems with classification as well as accuracy. Hence, the data mining algorithm called “Fuzzy k-means” method is employed to analyze and examine the ADHD projected here. It involves two steps – preprocessing and classification. First the given data are preprocessed to eradicate the noisy, redundant and incompatible data. Hence, the preprocessing is executed by Support Vector Machine (SVM). After preprocessing, the samples of ADHD are again classified and it is divided into moderate ADHD (ADHDmod) and high ADHD (ADHDhigh) sets. When integrating the SVM with Fuzzy k-means clustering algorithm, the performance measures are evaluated in terms of accuracy, specificity and

sensitivity. The results show that the projected work offers a better accuracy rate when comparing it with K – means clustering algorithm.

**Keywords:** *ADHD, K – Means, Fuzzy K – Means, Support Vector Machine*

### 1. INTRODUCTION

The problem is mentally and has a devastating effect on young children, especially at school. The neural disorders can be studied with the help of the volume of brain. The children who have reduced volume of brain are affected with ADHD. The category is based on children as well as adolescents. The symptoms will occur in both the categories, and must be analyzed to find out the area of affected features, whether temporal and frontal grey features [4]. The main theme of this work is to determine and classify the ADHD patients. The ADHD patients were categorized as medium ADHD and high ADHD. The vital part called as “caudate nucleus” helps in learning, cognition, memory as well as planning. The functionality may be changed due to the damages in brain or any abnormal situations which indicate the existence of ADHD. The volume of caudate nucleus dissimilarities results in brain abnormality. Since the occurrence varies the differences cannot be measured quickly. Hence, the development of Support Vector Machine (SVM) plays a major part in preprocessing the data points resulting in the removal of noise. Then, Fuzzy K – Means clustering method is considered for measuring the data points based on their similarities to determine the moderate ADHD and high ADHD. Final output was compared with current K – Means

---

<sup>1</sup> *Research Scholar, Dept of CA, CS & IT, Karpagam Academy of Higher Education, Coimbatore, India.*

<sup>2</sup> *Professor, Dept. of CA, CS & IT, Karpagam Academy of Higher Education, Coimbatore, India.*

clustering and it is proved that the projected method is better than the previous methods.

## 2. RELATED WORK

Several studies are based on the changes in the structure of brain volume. The attribute called “gender” plays a crucial role in changes in the volume of the brain. The diagnosis method was proposed by Zanaty et al.[3] and the analysis included changes in the volume of the grey matter. Another technique was found by Xu et al.[2] and it focused on the volume changes in the right frontal lobe which is located in the middle part of the right temporal and the left side of cerebellum and basal ganglia. The brain volume changes were proposed by Cherubini et al.[1],and the changes in parts like thalamus, caudate nucleus and putamen were analyzed.

## 3. EXISTING SYSTEM

### 3.1 K–MEANS CLUSTERING

The K – Means clustering is a technique for collecting data points and they are grouped together with similarities. The algorithm works as follows:

- a) Initialize k number of clusters
- b) k denotes the centroids
- c) After determining the centroid, the data is assigned to the nearest cluster
- d) Centroids are kept as small as possible
- e) Means refers to identifying the centroids
- f) Number of iterations is defined successfully.

The ADHD samples are analyzed and k numbers of clusters are assigned to the data points. Then the samples are classified based on the data points associated with the same clusters. But the drawback is that the prediction of k value becomes complicated [6]. The technique fails to run globally. And also the diverse segregation may end in diverse cluster outcomes.

## 4. MATERIALS AND METHODS

### 4.1 Support Vector Machine

ADHD is a disorder in behaviour. The behaviour patterns are deeply analyzed by considering factors

like lack of attention and hyperactivity [5]. These types of features cannot be seen among the children all the time and they may vary. Hence, the task of diagnosis becomes very tedious. To make the process easy, artificial-intelligence-based technique is employed for disorder diagnosis. The technique called as “Support Vector Machine” is used to analyze the disease. The difficulty in the intricate diagnosis can be overcome by this method. The SVM is utilized for preprocessing.

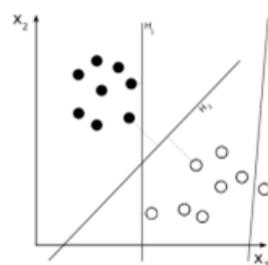


Figure No: 1 Support Vector Machine

The supervised learning algorithm is SVM and it helps in classification. The construction deals with various hyper planes used for classifying the tasks. The data separation can be highly improved by hyper plane method. The hyper plane is represented in figure 1.

The algorithm works thus:

- a) The data are collected and stored in a spreadsheet.
- b) Next step is the important step in preprocessing the information. SVM is effectively used for data preprocessing to remove the noisy data.
- c) The diagnosis is made with the help of SVM.
- d) Then the process is integrated with Fuzzy K – means clustering to provide good accuracy.

### 4.2 FUZZY K – MEANS CLUSTERING

After removing the noisy data, the method is combined with Fuzzy K – means clustering to classify the moderate ADHD and High ADHD among the children. This clustering is the most efficient technique. It helps to determine both the hard and soft clusters depending upon their probability. The k

number of clusters is initialized after preprocessing.

The steps are illustrated as follows:

- a) Initialize the k number of clusters
- b) The probability computation is performed which is in the same cluster
- c) Run the iteration – Calculate m value. When the value is greater than or equal to 2, then this is considered to be the closest center point of the cluster
- d) Clusters are made to run and hence the classification of moderate ADHD and high ADHD is predicted
- e) Distance measure is used to calculate the nearest point to the clusters.

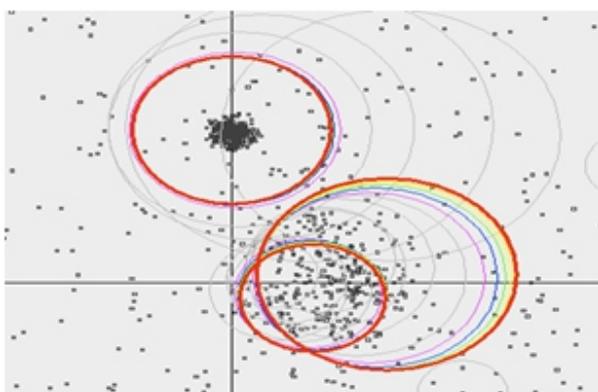


Figure No: 2 ADHD Classification

The fuzzy K – Means is an iterative method and the clusters are shown in different colors like yellow, blue, orange, red, violet and blue and is shown in figure 2. It is shown that red color indicates children who have High ADHD and blue those who have moderate ADHD. The classification accuracy is high when compared to other methods.

**5. RESULTS AND DISCUSSION**

The projected work is utilized to classify the disorder based on two things: a) Moderate ADHD and b) High ADHD. The data preprocessing is done by SVM classifier to minimize the noise. Then the data are classified by Fuzzy K – means clustering technique. The classification results are provided in table 1 and figure 3 and the accuracy results in table 2 and figure 4.

Table No: 1 Classification Rate

Classification Rate	Methods	
	K – Means	SVM + Fuzzy K - means
Moderate ADHD	82%	87%
High ADHD	91%	96%

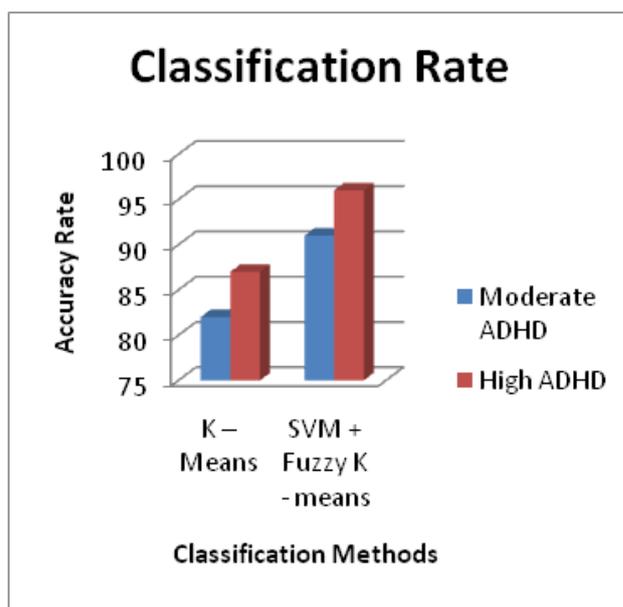
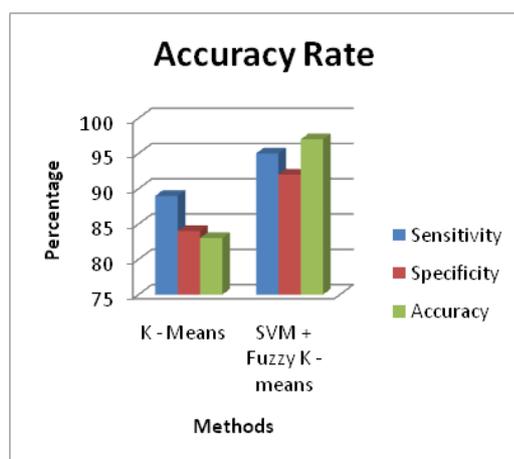


Figure No: 3 Classification Rate

The ADHD among the children of different ages is considered and the method is evaluated based on the SVM and Fuzzy K – means algorithm. The results are more accurate. The accuracy is about 92.3% when it is diagnosed by integrating SVM and Fuzzy K – means algorithm for ADHD analysis.

Table No: 2 Accuracy Rate

Methods	Sensitivity	Specificity	Accuracy
K-Means	89	84	83
SVM + Fuzzy K-means	95	92	97



**Figure No: 4 Accuracy Rate**

The support vector machine affords good interpretation results. The diagnosis theory is very efficient and consumption of time is less. The method tells us that the overall accuracy of existing method is 83% and the specificity rate and sensitivity rate are predicted to be 89% and 84% respectively. But the proposed system offers good results with the accuracy rate of 97% and the specificity rate and sensitivity rate are predicted to be 95% and 92% respectively.

## 6. CONCLUSION

In the present world, a greater number of children are affected by ADHD problems. Mental disorders are faced by more children. The method is focused on diagnosing the various disorders by various classification algorithms. The work helps to remove the noisy data by Support Vector Machine algorithm and after preprocessing the classification is performed by assimilation SVM and Fuzzy K – means clustering algorithm. The clustering helps to group similar data and hence the fuzzy k – means classifies similar results and enhances the accuracy results. The method classifies the samples based on moderate ADHD and High ADHD. The projected work provides good performance results.

## 7. REFERENCES

- [1] Cherubini A, Péran P, Caltagirone C, Sabatini U, Spalletta G. Maturing of subcortical cores: Microstructural, mineralization and decay adjustments estimated in vivo utilizing MRI. *NeuroImage*. 2009.
- [2] Xu J, Kobayashi S, Yamaguchi S, "Gender orientation consequences for age related changes in mind structure", *AJNR Am J Neuroradiol*, 2000.
- [3] Zany E.A, "Assurance of Gray Matter (GM) and White Matter (WM) Volume in Brain Magnetic Resonance Images (MRI)", *International Journal of Computer Applications (0975 – 8887) Volume 45–No.3, May 2012*.
- [4] Rashedur M. Rahman and Farhana Afroz, "Correlation of Various Classification Techniques utilizing Different Data Mining Tools for Diabetes Diagnosis", *Journal of Software Engineering and Applications*, Vol. 6, No. 3, pp. 85-97, 2013.
- [5] David D. Luxton, "AI in Psychological Practice: Current and Future Applications and Implications", *Professional Psychology: Research and Practice*, Vol. 45, No. 5, pp. 332-339, 2014.
- [6] Jerzy Gomuła, Krzysztof Pancerz and Jarosław Szkoła, "Classification of MMPI Profiles of Patients with Mental Disorders-Experiments with Attribute Reduction and Extension", *Proceedings of International Conference on Rough Sets and Knowledge Technology*, pp. 411-418, 2010.