A Survey on Mining Algorithms for Predictive Analytics of Asd

K. Vijayalakshmi **Dr. M . Vinayakamurthy** **Dr. V. Anuradha** **Pavan Chunduri**

ABSTRACT

Autism is a lifelong developmental deficit that affects how people perceive the world and interact with each others. An estimated one in more than 100 people has autism. The Autism disorder is most commonly seen in boys than girls. The commonly used tools to analyzing the dataset of autism are FMRI, EEG, and more recently "eye tracking". A several tests has been made on the ones who are suffering from Autism which helps to study on ASD and provide with the results to take further steps to treat them. In this research paper we study the algorithms which improves the prediction levels of Autism and provide a review on which algorithm is best suited to predict the autism in early stages. The data mining algorithms which used to predict the autism were mainly categorized in 4 types.

Keywords: **Autism, FMRI, ECG, EYE Tracking, PCA, Data Mining**

I. INTRODUCTION

Autism Spectrum Disorder (ASD) which came into the light very recently and became very pressing matter in today world. The Autism which mainly effect the children than the adults and because of that it became for the children's very hard to communicate with the others. The people who are suffering from ASD doesn't able to live a quiet life as it effects their daily needs and unable to finish very easy tasks due to the Autism, to find out the people who are suffering from Autism and provide with better treatment to cure the Autism at the very early stages. In the past there were several researches has been made to find the ASD in the very early stages in children and find out the kind of treatment to be made.

There were many different techniques in different concepts such as Data mining, Machine Learning, Deep learning to predict the ASD in the very early stages in different age people. In the past 10 Years different people were took the treatment for the ASD has increased gradually. Exactly after 5 years with the help of the previous data a research has been did and came to acknowledge that the treatment cost to ASD has increased to 5 times their yearly expenses, the major part of the money goes just to find out the person is suffering from ASD or not .The Autism Spectrum Disorder (ASD) can be predicted by different algorithms which gives the idea to cure it at the early stages.

II. LITERATURE SURVEY

The Autism is mainly effects the human mind and unable to communicate and co-exist with the other people and eventually it became harder to do the basic work. In the recent study it is made clear that to cure the Autism and cost just for the treatment is 5 times the yearly income of a common man. By using the different data mining algorithms it is easy to predict the ASD at the very early stage [1]. At the early stages when you known the ASD in children you can take precautions and able to prevent it and minimize the effect when they are growing. But at this point to get the data with the accurate results from doctors is highly impossible [2]-[4]. In the paper [7] different data mining algorithms is used to predict the ASD (Autism Spectrum Disorder).

In this paper we used different data mining algorithms to predict the Autism disorder. In the work of [3] the two algorithms LDA and KNN which provide the prominent values which helps to find the better algorithm to find the autism disorder and help to take precautions in the early stages. With the help of SVM (Support Vector Machine) in [5] finding the ASD is a bit tough when you are predicting the ASD, at first you need to clearly understand the three functions which are: linear kernel, polynomial kernel and radial core.

III. DATA MINING ALGORITHMS

a. K-Means Algorithm:

The K-Means algorithms used to evaluate the ASD through numeric values which obtained from the previous results. Where the other data mining algorithms used to the data sets which they obtained from the previous results and the data set may be numerical or other sources. The past methods which used to find the ASD were studied and provided with the results. In the Algorithms they take the numerical values from the previous results and check with the current results and calculate the disorder. In the K-Means algorithm we used a method to predict the Autism Spectrum Disorder (ASD) with the different data sets.

A period time of work is invested in this to define the work with the Autism. With the help of the K-Means algorithm to predict the ASD the sample has taken from the school children who are suffering from ASD. With the help of the algorithm we can predict that which school age the ASD is common and also with the help of the data the precautions can be taken and with the outcome of the results we can decide on how to cure it. The ASD can be examined with the help of different test and we come to an analysis that which age it affects more. With the Data mining algorithms we can come to an conclusion to predict the Autism Spectrum Disorder (ASD) at the early stages.

With the help of the previous study made on algorithms that take on the numerical values to predict the AD in the children that helps to take measures. With the recorded samples used to provide the better results on the ASD to Which helps to take the measures to cure the ASD. The results or samples can be obtained with the help of tests which can help to obtain the sample results and the data can be used to provide the better results on finding the ASD.

K- Nearest Neighbor Algorithm

The other Data mining algorithm which helps to predict the ASD is K-Nearest algorithm. The K-Nearest algorithm which provides the results majorly with the help of computers. But the early time the computers didn't have the major role, due the results will be obtained through the old fashioned way. Of all the data mining algorithms the KNN algorithm is considered the most understandable algorithm compared to the other algorithms. After the computers has made to the world the KNN algorithm has made easy to predict the ASD. With the previous data sets we can put the numerical values to the KNN algorithm.

The KNN used different methods the main commonly used method is the Euclidian. In this algorithm there are different steps to be taken, at the very first step with the help of the data acquired we will calculate the taken data. After the data is taken it will be calculated. In the next step the values will be taken and provided it to the class. Due to which the k- value always determines with 1.

Linear Discriminant Analysis

Within the Data mining algorithms the LDA algorithm is mostly used algorithm that which help to predict the ASD. The Linear Discriminant Analysis which used to take the sample data sets and calculate the same data sets to provide the better results. In this algorithm to calculate the data certain equations is used to get the results to predict the Autism Spectrum Disorder (ASD). The LDA algorithm that which it provides the accurate results with the help of the data sets.

The Autism which mainly affects the children than the adults and because of that it became for the children's very hard to communicate with the others. Without the proper communication with the others at the early stage will lead to effect in future. The predicting and treating the ASD at the early stages is very important. For a sample we have taken the data set of 300 datasets. In curing the ASD the LDA and KNN algorithm were taken for the testing of the sample data set. Due to which the k- value always determines with 1.

In the table 1 we have taken the sample data sets and with the help of LDA and KNN algorithm to find out the True and false on Autism disorder. With the true it indicates that the Autism is predicted on the children and if the false indicates that the Autism didn't predict correctly in children and we will get a accurate results by performing again. He values which taken for the algorithms LNN and KNN is given in the below table.

Table 1: Number of Observations

	LDA	KNN
TP	80	81
FP	5	4
TN	79	76
FN	25	28

In the above table the data set which we taken almost 65-70 percent of the data is taken as the training set and almost 35-30 percent of the actual and real data set which is taken from the previous data set. With the help of both the algorithms the results were taken and observed the data set where it predicts the Autism disorder at the early stages. Out of the two algorithms the LDA provides with the better and accurate results than the KNN algorithm.

b. Support Vector Machine (SVM)

Another data mining algorithm thst helps to predict the Autism disorder is Support Vector Machine (SVM). The only drawback in the SVM algorithm is that it is advanced and kind of hard to work to this with the beginners and it won't provide any better results compare to other algorithms, but with the help

of the certain steps it can be avoided. Support Vector Machines classification method is a supervised classification technique, the main disadvantage of which is the choice of the kernel function. In this Research, briefly reflect apply the three functions which are: linear kernel, polynomial kernel and radial core in order to clearly understand which is the most suitable for the data of the adults as well as those of the children.In this method, automatic learning is taken as half of the autistic people and half of the normal, The nucleus chosen at the beginning is linear then polynomial finally a nucleus Radial base.

IV. CONCLUSION

The paper work which we provide the work of the different data mining algorithms. In the study, we predict the Autism disorder with the help of the different algorithms. A sample data set is used in different algorithms. In the different algorithms the LDA and KNN where much used and gives a close to the finding the levels of the ASD. With the study of the different algorithms we came to the conclusion that which algorithm that provides the better results in predicting the Autism Spectrum Disorder (ASD). With study of the algorithms we came to the conclusion that the LDA provides the accurate results compare to the other algorithms. In all the algorithms that provide the better result the LDA has more accuracy compare to the other algorithms like K-Means. KNN , SVM and it has more that one percent higher compare to the other algorithms.

V. REFERENCES

[1] J. Fan et al., "A Step Towards EEG-based Brain Computer Interface for Autism Intervention," presented at the 37th Annual International Conference of the IEEE Engineering in Medicine and Biological Society (EMBS), in press, 2015.

[2] O. Grynszpan, J.-C. Martin, and J. Nadel, "Multimedia interfaces for users with high functioning autism: An empirical investigation," International Journal of Human-Computer Studies, vol. 66, no. 8, pp. 628-639, 2008.

[3] D. L. Christensen, "Prevalence and characteristics of autism spectrum disorder among children aged 8 years—autism and developmental disabilities monitoring network, 11 sites, United States, 2012," MMWR. Surveillance Summaries, vol. 65, 2016.

[4] M. Uljarevic and A. Hamilton, "Recognition of emotions in autism: a formal meta-analysis," Journal of autism and developmental disorders, vol. 43, no. 7, pp. 1517-1526, 2013.

[5] Nermend, Kesra. "Vector calculus in regional development analysis: Comparative regional analysis using the example of Poland". In Springer Science & Business Media, 2009.

[6] Adankon, Mathias M., and Mohamed Cheriet. "Support vector machine." In Encyclopedia of biometrics, Springer US, 2009, pp 1303-1308.

[7] Y. Liu, Z. Li, H. Xiong, X. Gao and J. Wu, "Understanding of internal clustering validation measures," in IEEE 10th International Conference on Data Mining (ICDM), 2010.

[8] Frank E, Hall M, Holmes G, Kirkby R, Pfahringer B, Witten IH, Trigg L (2009) Weka-a machine learning workbench for data mining. In: Data mining and knowledge discovery handbook. Springer, Boston, pp 1269–1277.