A Novel Approach to Predict the Learning Skills of Autistic Children using SVM and Decision Tree

M.S. Mythili[1], A.R.Mohamed Shanavas[2]

[1]Ph.D Research Scholar, Bharathidasan University & Assistant Professor,
Department of Computer Applications, Bishop Heber College, Tiruchirappalli 620 017, TamilNadu, India.
[2]Associate Professor, Dept. of Computer Science, Jamal Mohamed College,
Tiruchirappalli 620 020, TamilNadu, India

Abstract - **The present trends in education are moving toward further comprehensive educational experiences for children with autism spectrum disorder, totally different learning styles, and differentiated instructional practices to support all children to succeed. The very important goal of the paper is to predict and forecast the autism spectrum disorder children performance by applying data mining classification algorithms in Weka tool. The data mining tool is accepted as a decision making tool to facilitate higher resource utilization in terms of autism children performance. The classification algorithms may be specifically mentioned as J48 and Support Vector Machine. The results of the classification model deals with the accuracy level, error rate level, confusion matrices, efficiency of classifiers and additionally the execution time. Sequential Minimal optimization (SMO) algorithmic rule with Normalized Polykernel used in performing classifier SVM and the J48 algorithmic rule is employed in constructing decision trees.**

Keywords- **Decision Tree, Support Vector Machine**

1. INTRODUCTION

Data mining is the step of the "Knowledge Discovery in Databases" or KDD associate with an interdisciplinary subfield of science. It is the computational process of discovering patterns in massive information sets involving strategies at the intersection of artificial intelligence, machine learning approach, statistical information systems. The general goal of the data mining method is to extract information from a data set and transform it into an understandable structure for further more use. Aside from the raw analysis step and rules, it discovers and involves in the information system aspects, processing, models, interestingness measures and the data visualization. The important particular data processing task and role is that the automatic or semi-automatic analysis of enormous quantities of knowledgeable data to extract antecedently unknown fascinating patterns such as groups of data records, uncommon records and dependencies. This typically involves information techniques such as spatial indices. The patterns will then be seen as a form of outline of the raw input data, and will be used as a vital thing in more analysis. The connected terms data dredging, data fishing, and data snooping check with the employment of data mining strategies to sample element parts of a massive population data set that are too tiny for reliable statistical inferences to be created concerning the validity of any patterns discovered. These strategies will, however, be employed in making new hypotheses to check against the larger data populations.

The application of data mining is widely used in educational system. Educational data mining is a rising field which can be effectively applied in the field of education. The knowledge that emerges can be used to perceive children reading skills, children writing skills, children language and communication skills, children handwriting skills, children spelling skills and the children memory skills. The data mining system is vital and crucial to deeply measure the autism children learning skills performance improvement. The classification algorithms are often wont to classify and analyze the autism children data set in perfect accurate manner. The main objective of the paper is to use data mining methodologies to forecast and predict the autism children performance. Data mining provides different types of tasks that could be used to study the autism children performance. In this paper, the classification task is utilized and employed to measure and gauge the autism children performance and deals with the accuracy, confusion matrices and also the execution time taken by the various classification data mining algorithms.

This paper is catalogued as follows. Section two describes Autism children role with their signs and symptoms. Section three gives the related work Section four presents the idea of Classification and discusses the aspects of classification algorithmic rule. Section five elaborates a Data Preprocessing. Section six explains the Implementation of model construction. Section seven describes the results and discussions. Section eight provides the conclusion.

II.AUTISM

Autism is a neurological disorder that have the impair communication, socialization and behavior. It is typically diagnosed within the first three years of life and it is commonly four times more common in boys than in girls. The types of Autism might not be diagnosed till years later once the child enters school, owing to late-occurring social deficits or the problem finding difficulty in playing with others. Although awareness and understanding have greatly enhanced over the past few decades, many people are still unaware of the actual affect of Autism. It will become an overshadowing factor considered in each side of life, as well as education, establishing and maintaining relationships, responding to pain and discomfort, and within the ability to specific feeling. Symptom severity in

Autism will vary from mild to severe. As parents reach the diagnosing, treatment and education stages of Autism, they will hear many alternative terms used to describe their child. They might embrace words such as autistic-like, non-verbal, developmental delay, autistic tendency, savant, high-function autism, and low-function autism. The necessary factor to appreciate is that each child with Autism is totally different. The diagnosing is, children with Autism are ready to learn, perform profitably in society and show positive gains with applicable education and treatment plans in place. Without appropriate support of the teachers and their parents, the child might never notice his full potential.

III. RELATED WORK

Prud'hommeaux et al. analyze the problems for classification of non standardized text of machine learning algorithms. Kathleen T Quach enumerates the issues of the classification problem which ASD has a terribly heterogeneous disorder which will have subgroups with drastically has completely different genetic expression signatures. To enhance and improve the classification algorithmic rule, it should be terribly helpful to stratify the ASD category into subgroups and enrich the input set with clinical measures. Alexander Genkin et al. explains a simple Bayesian logistic regression approach that uses a Laplace prior to avoid over fitting and produces sparse predictive models for text data. They applied this approach to a spread of document classification issues and show that it produces compact predictive models a minimum of as effective as those made by support vector machine classifiers or ridge logistic regression combined with feature choice selection.

IV. CLASSIFICATION

This method is incredibly helpful and correct to category data into predefined categorical class labels. Classification could also be a two step technique consisting of training data set and therefore the testing data set. In the first step, a model is made by analyzing the data tuples from the training data having an attributes. For each tuple within the training data, the value of class label attribute is clearly understood. Classification rule is applied on training data to create the model method. In the second step of classification, the test data is utilized to look the clear accuracy of the model method. If the accuracy of the model is acceptable then the model may be utilized to classify the unknown data tuples. The basic algorithms utilized in the paper for classification are decision tree classifier and support vector machine.

A. Classification Algorithms

This analysis research paper contains a common decision tree classifier C4.5 (J48) and a Kernel method (SVM).The classifiers are mentioned briefly.

B. Decision tree classifiers

A decision tree is a flow chart resembling a tree structure, wherever each internal node is notated by rectangles and therefore the leaf nodes are notated by ovals. This algorithmic program is commonly used as a result of the implementation is simple and easier to grasp compared to the other different classification algorithms. Decision tree starts with a root node that allows the users to take needed actions. From this node, users split each node recursively according to decision tree learning algorithmic program. The final ultimate result is a decision tree in which every branch associate an outcome.

C. J48

This algorithmic program may be a successor to ID3 developed by Quinlan Ross. It's in addition supported the Hunt's algorithmic program.J48 handles every categorical and continuous attributes to form a decision tree, so as to handle continuous attributes. J48 splits the attribute values into 2 partitions supported on the chosen threshold. It additionally handles missing attribute values. J48 has the idea of Gain Ratio as an attribute selection measure to form a decision tree. It fully and clearly prunes the biasness of information gain once there are several outcome values of an attribute. At the first level, calculation of the gain ratio of each attribute is needed. The root nodes are the attribute whose gain ratio is maximum. J48 uses pessimistic pruning to get induce of unessential branches within the decision tree to reinforce the accuracy of classification.

D. Support Vector Machine

In machine learning, support vector machine are supervised learning method with associated learning algorithms. It analyzes the data and acknowledges patterns. It is primarily used for classification and multivariate analysis. An SVM training algorithmic program builds a model that assigns new samples into one class or the other, using it as the non-probabilistic binary linear classifier. An SVM model could be an illustration of the examples as points in area, mapped so that the samples of the separate classes are divided by a transparent gap that's as wide as attainable. New examples are then mapped into that very same area and expected to belong to a class based on which side of the gap they fall on. In addition to acting linear classification, SVM efficiently perform a non-linear classification is called the kernel trick, implicitly mapping their inputs into high-dimensional feature area.

V. DATA PREPROCESSING

Datasets utilized within the classification algorithmic rule ought to be clear, accurate and can be preprocessed for handling missing or redundant attributes. The data is to be handled with efficiency to induce the best outcome from the Data Mining process.

A. Attribute Identification

Dataset collected from children database consists of

Attributes	Description
Atten	Attention Skills
Hand writ	Hand Writing Skills
Spelling	Spelling Skills
Language	Language Skills
Read	Reading Skills
Write	Writing Skills
Memory	Memory Skills
Child	Levels of child autism

VI. IMPLEMENTATION OF MODEL CONSTRUCTION

Weka is open source software system that implements a large collection of machine learning algorithms and is widely utilized in data mining applications. From the above data, student.arff file was created. This file was loaded into a WEKA explorer. The autism children learning skills is influenced by various factors like attention Handwriting ,spelling,language,Read,write,memory skills and result from the different children.40 samples were taken for the implementation. The classify panel permits the user to use classification algorithms to the dataset, to estimate the accuracy of the resulting predictive model, and to visualize the model.

The decision tree classifier (J48), Normalized PolyKernel based classifier (SVM) were enforced in weka. Under the "Test options", the 10 fold cross validation is chosen.

VII. RESULTS AND DISCUSSION

The analysis and interpretation of classification is time consuming process that needs a deep understanding of statistics. The process needs a large amount of time to finish and expert analysis to look at the classification and relationships within the data.
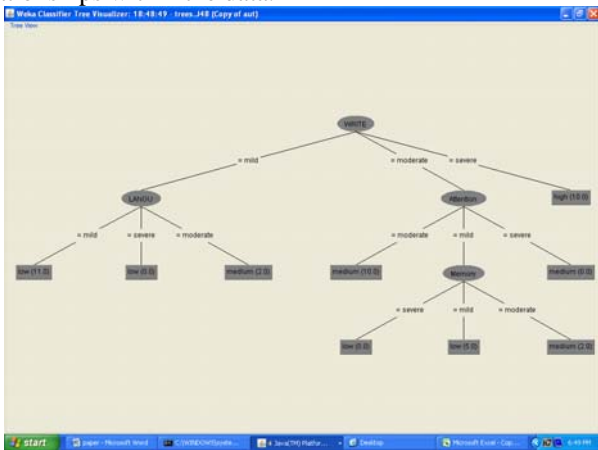

Fig 1 Visual image of generated decision trees

The Writing Skills of the autism children' is taken as root node from that Language Skills, Attention Skills and with the Memory Skills and so on. The knowledge represented by decision tree can be extracted within the form of IF-THEN rules.

1. IF write="mild" AND langu="mild" THEN child="low"
2. IF write="mild" AND langu="severe" THEN child="low"
3. IF write="mild" AND langu="moderate" THEN child="medium"
4. IF write="moderate" AND attention="moderate" THEN child="medium"
5. IF write="severe" THEN child="high"
6. IF write="moderate" AND attention="mild" AND memory=severe THEN child="low"
7. IF write="moderate" AND attention="mild" AND memory=mild THEN child="low"
8. IF write="moderate" AND attention="mild" AND memory=moderate THEN child="medium"
9. IF write="moderate" AND attention="severe" AND THEN child="medium"

From the above set of rules an inescapable conclusion emerges the writing skills is considerably related with student performance. From the rule set it was found that communication, hand writing, spelling, language, reading, writing, memory, and the levels of autism children are the different factors are of high potential variable that have an effect on autism children performance.

TABLE II
PERFORMANCE RESULT OF CLASSIFIERS

Evaluation Criteria	J48	Support Vector Machine
Timing to build the model(sec)	0.02	0.28
Correctly classified instances	36	38
Incorrectly classified instances	4	2
Accuracy	90 %	95%

In Table II the percentage of correctly classified instances is usually referred as accuracy of the model. Hence Support Vector Machine with Normalized Polykernel can be termed as more accurate than other classifiers.
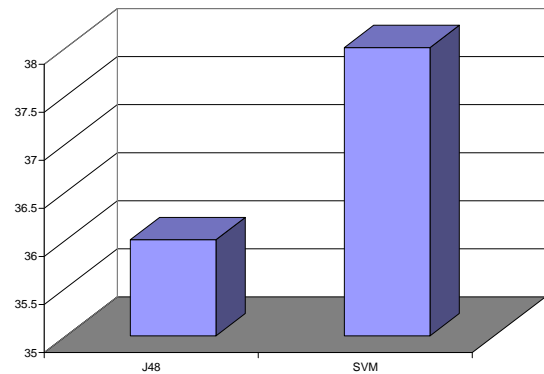

Fig 1 Accuracy of the classifier algorithm

The Figure 1 shows that the graphical representation of accuracy results of autism children learning skills based on dataset. It clearly reveals that Support Vector Machine is a very best classifier for analyzing the autism children performance result consuming less time coupled with good accuracy.
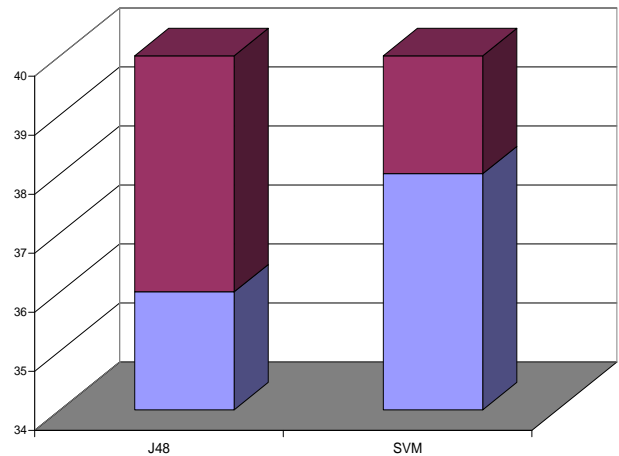

Fig 2 EFFICIENCY OF DIFFERENT CLASSIFIERS

The Figure 2 explains the graphical representation of correctly classified instances of results of autism children learning skills based mostly on autism children dataset. The highest percentage of correctly classified instances is the Support Vector Machine classifier.

TABLE III
ERROR MEASUREMENT FOR CLASSIFIERS

Evaluation Criteria	J48	Support Vector Machine
Kappa statistic	0.8462	0.9234
Mean absolute error	0.0625	0.0233
Root mean squared error(RMSE)	0.2254	0.2154

In Table III Kappa statistics is a measure of the degree of non random agreement between observers and measurement of a particular categorical variable. The root mean square error and Mean absolute error of Support Vector Machine are minimum when compared to other classifiers. Therefore the Support Vector Machine is that the efficient classification technique among remaining classification technique.
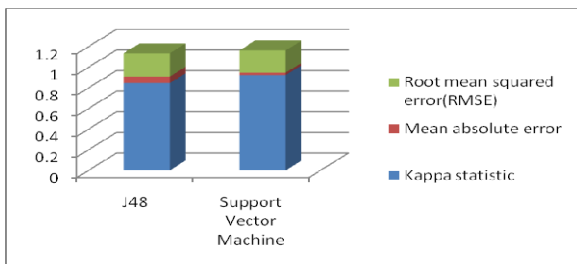

Fig 3 ERROR RATE OF DIFFERENT CLASSIFIERS

The Figure 3 compares errors among completely different classifiers (root mean square error and Mean absolute error) Support Vector Machine has lower error rate compared to different classifiers. Therefore the Support Vector Machine is the efficient classification technique among remaining classifiers.

TABLE IV
CLASS LABEL ACCURACY FOR CLASSIFIERS

classifier	TP	FP	Precision	Recall	Class
J48	1	0.167	0.8	1	Low
	1	0	1	1	High
	0.714	0	1	0.714	Medium
Support Vector Machine	1	0.083	0.889	1	Low
	1	0	1	1	High
	0.857	0	1	0.857	Medium

The Table 4 clearly shows the performance of every classifier based on the true positive rate (TP rate) and false positive rate (FP rate), precision, recall and different measures. These measures are very helpful for comparing the classifiers based on the accuracy. The Support Vector Machine algorithm outperforms all different classifiers within the children dataset.

TABLE V
CONFUSION MATRIX

Classifier	High	Low	Medium	Class
J48	0	16	0	Low
	10	0	0	High
	0	4	10	Medium
Support Vector Machine	0	16	0	Low
	10	0	0	High
	0	2	12	Medium

The Table 5 reveals that the confusion matrices are very helpful for analyzing the classifiers.

VII. CONCLUSION

The work explores the potency of machine learning algorithms in deciding the influence of result, various factors like attention Hand writing, spelling, language, Read, write, memory skills and result from the autism school student learning skills and analyze of autism children performance. It is discovered that Support Vector Machine is best than that of different algorithms employed in the study. This study is going to be terribly useful for the educational institutions. In future, it is doable to increase the analysis by using different classification techniques and association rule mining for the autism children dataset.

REFERENCES

[1] "Guidelines for Educating Students with Autism Spectrum Disorders", October 2010 Virginia Department of Education, Office of Special Education and Student Services.
[2] "Guidelines for Identification and Education of Children and Youth with Autism ", July 2005, CONNECTICUT STATE DEPARTMENT OF EDUCATION Division of Teaching and Learning Programs and Services Bureau of Special Education
[3] Felix D. C. C. Beacher et al., "Sex Differences and Autism: Brain Function during Verbal Fluency and Mental Rotation", June 2012 | Volume 7 | Issue 6 | e38355
[4] Volkmar FR, Klin A (2000) Pervasive Developmental Disorders. In: Kaplan and Sadock's Comprehensive Textbook of Psychiatry CDROM 7th edition. Publisher: Lippincott Williams and Wilkins, Philadelphia, PA.
[5] "Autism Spectrum Disorders" Centre for Developmental Disability Health Victoria Building 1, 270 Ferntree Gully Road, NOTTING HILL, VIC 3168 Telephone: (03) 9902 4467 Facsimile: (03) 8575 2270 E-mail: cddh@monash.edu.
[6] "Autism Spectrum Disorders: Information Pack" Autism Victoria Inc. ABN 15 600 724 949 A14601C 24 Drummond St, Carlton, 3053, Postal Address: P.O. Box 374, Carlton South, 3053 E: info@autismvictoria.org.au T: 03 9657 1600
[7] Baron-Cohen S (2009) Autism: the empathizing-systemizing (E-S) theory. Annals of the New York Academy of Sciences. Mar;1156: 68–80.
[8] Eric Zander et al., "An introduction to autism", AUTISMFORUM Handikapp & Habilitering, Box 17519, 118 91 Stockholm 08-690 60 52, www.autismforum.se, autismforum@sll.se Zander, E. An introduction to autism ¼ 1/9 2004.
[9] Kathleen T Quach et al., "Application of neural networks in classification of autism diagnosis based on gene expression signatures"
[10] Demuth et al. "Neural network toolbox for use with MATLAB.",1993. Rachna Ahuja et al, / (IJCSIT) International Journal of Computer Science and Information Technologies, Vol. 5 (2) , 2014, 2166-2170 www.ijcsit.com 2169
[11] Nagesh Adluru et al., "Characterizing brain connectivity using - radial nodes: application to autism classification". [14] Hagmann, P., Kurant, M., Gigandet, X., et. al.: Mapping human whole-brain structural networks with Diffusion MRI. PLoS one 2(7) (2007).
[12] Prud'hommeaux et al,"Classification of atypical language in autism", in Proceedings of the 2nd Workshop on Cognitive Modeling and Computational Linguistics, pp: 88-96, 2011.