IJISET - International Journal of Innovative Science, Engineering & Technology, Vol. 1 Issue 8, October 2014.
www.ijiset.com
ISSN 2348 – 7968

Diagnosis of Learning Disabilities in school going children using Data Mining Techniques: a survey

[1]**Margaret Mary**. T , [2]**Hanumathappa. M**

[1]Lecturer in MCA Dept., Sambhram Academy of Management Studies, Bangalore-97 and Research Scholar, Dept., of CS, Bharathiar University, Coimbatore-641 046

[2]Professor, Department of Computer Science & Applications, Bangalore University, Bangalore-56

Abstract

This paper is a survey paper on the prediction of learning disabilities (LD) in school- age children. LD is a Disruptive Behaviour Disorder characterized by the presence of a set of chronic and impairing behaviour patterns that display abnormal levels of inattention, hyperactivity, or their combination.[3] Since most individuals especially children display these behaviours from time to time, it is be difficult to differentiate behaviours that reflect LD from those that are a normal part of growing up which makes the diagnosis a tricky job.

In this paper, surveys on implementation methods are using difference well know artificial intelligence techniques like Support Vector Machine, Neural Network and Decision tree. SVM algorithm for the diagnosis of the disorder. The major advantage of using SVM is that it helps in controlling the complexity of the problem of diagnosing. Neural network method is used for classification, clustering, feature mining, prediction and pattern recognition. Decision tree are powerful and popular tool for classification and prediction [3]. These techniques are emphasis on application of data mining, and playing vital role for classification, analysis and solving the design of an LD prediction tool based on machine learning technique. To implement these classification techniques, required different source and methods of data collection, the data set, data distribution, normalization and attribute normalization. The main aim of this paper is to present a survey of all of researchers who worked in this area and explaining different approaches to identify learning disabilities in school going children.

Index Terms - Preprocessing, Learning Disability, Neural Network, Support vector machine and Decision tree.

1. Introduction:

Data Mining is a general term, which describes a number of techniques used to identify pieces of information or decision-making knowledge in data. A common misconception is that it involves passing huge amounts of data through intelligent technologies that, alone, find patterns and give magical solutions to business problems. Data Mining is an interactive and iterative process. Business expertise must be used jointly with advanced technologies to identify underlying relationships and features in the data. A seemingly useless pattern in data discovered by Data Mining technology can often be transformed into a valuable piece of actionable information using business experience and expertise.

Many of the techniques used in Data Mining are referred to as “machine learning” or “modeling”. Historical data are used to generate models, which can be applied at a later date to areas such as prediction, forecasting, estimation and decision support [20].

1.1. Preprocessing

Data mining is an essential step in the process of knowledge discovery. It consists of an iterative sequence of steps namely data cleaning, data integration, data reduction, data transformation, data mining, pattern evaluation and knowledge presentation. Data preprocessing is the pre stage of data mining. The data is ready for the mining process only after completion of data preprocessing,[3] and classification is a form of data analysis that extracts models describing important data classes. It consists of two step process. In the first step, a classification model based on previous data is build. The model describes a predetermined set of data classes or concepts [1]. This is the learning step or training phase and in the second step, the accuracy of classifier is determined by using the model to classify new data. The data mining or knowledge discovery process is shown in Figure- 1

IJISET - International Journal of Innovative Science, Engineering & Technology, Vol. 1 Issue 8, October 2014.
www.ijiset.com
ISSN 2348 – 7968

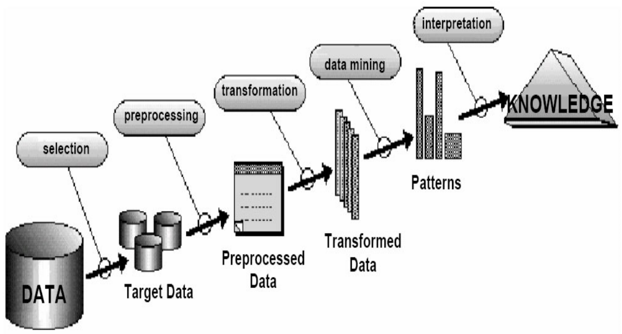

Figure- 1 Data Mining process

To analysis or preprocess the learning disabilities in children's must required data collection and data sets has different checklists containing general signs and symptoms or attributes related to LD are used for the general assessment of learning disability, where such LD assessment studies are generally conducting, are selected and with the help of Doctors and other Professionals / Resource Persons engaged there, the various LD assessment methods are studied and after subsequent evaluation with the help of these professional and from the experience gained, 16 prominent attributes are selected as below shown table-1

Table -1 list of Attributes

Sl. No.	Attribute	Signs & Symptoms of LD
1	DR	Difficulty with Reading
2	DS	Difficulty with Spelling
3	DH	Difficulty with Handwriting
4	DWE	Difficulty with Written Expression
5	DBA	Difficulty with Basic Arithmetic skills
6	DHA	Difficulty with Higher Arithmetic skills
7	DA	Difficulty with Attention
8	ED	Easily Distracted
9	DM	Difficulty with Memory
10	LM	Lack of Motivation
11	DSS	Difficulty with Study Skills
12	DNS	Does Not like School
13	DLL	Difficulty in Learning a Language
14	DLS	Difficulty in Learning a Subject
15	STL	Slow To Learn
16	RG	Repeated a Grade

1.2 Data Distribution - The percentage distribution of the 16 attributes present in the data set is considered by the various researchers shown in Figure below. Among the cases in the dataset, 72% cases (children) are having LD (LD – True) and the remaining 28% cases (children) have not LD (LD – False) as represented.

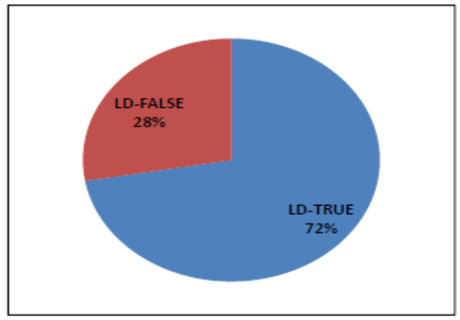

Figure 2 Distribution of LD cases in Data set

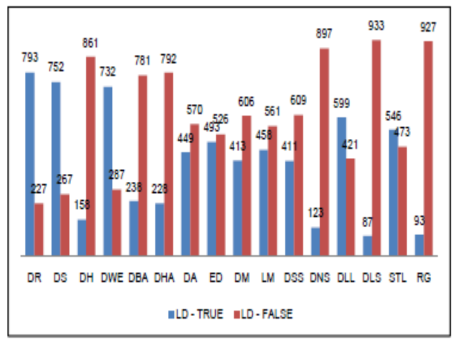

Figure 3 Significane of attributes in LD prediction.

1.3 Data Normalization - Data usually collected from multiple resources and stored in data warehouse. Resources may include multiple databases, data cubes, or flat files, during integration of data that we wish to have for mining and discovery, so data integration must be done carefully to avoid redundancy and inconsistency that in turn improve the accuracy and speed up the mining process[4] Data is transformed or consolidated into forms appropriate for mining. Data transformation involves data smoothing, data aggregation, data generalization, data normalization and data attribute construction or feature construction.

1.4 Attribute normalization - Normalization is particularly useful for classification algorithms involving neural networks, or distance measurements such as nearest-neighbor classification and clustering. Min-max normalization performs a linear transformation on the original data. Min-max normalization preserves the relationships among the original data values. An attribute A are normalized based on the mean and standard deviation of A. This method of normalization is useful when the actual minimum and maximum of attribute A are unknown. The number of decimal points moved depends on the maximum absolute value of A. Min Max Normalization transforms a value A to B which fits in the range[C, D]. The attribute data is scaled to fit in a specific range.

ISSN 2348 – 7968

$$B = \left\{ \frac{(A - \text{minimum value of } A)}{(\text{maximum value of } A - \text{minimum value of } A)} \right\} * (D - C) + C$$

The main method used is the min-max normalization. In data normalization, the attribute data are scaled so as to fall within a small specified range, such as -1.0 to 1.0 or 0 to 1.0 [5]

1.5 Learning Disabilities.

Learning disabilities, or learning disorders, are an umbrella term for a wide variety of learning problems. A learning disability is not a problem with intelligence or motivation. Kids with learning disabilities aren't lazy or dumb. In fact, most are just as smart as everyone else. Their brains are simply wired differently. This difference affects how they receive and process information.

Children and adults with learning disabilities vision, hear, and understand things differently. This can lead to trouble with learning new information and skills, and putting them to use. The most common types of learning disabilities involve problems with reading, writing, math, reasoning, listening, and speaking.[19]

2. Literature Survey on Implementation Methods

Julie M David et al [3] proposed implementation on classification method such as neural network machine, which are performed in weka. Neural networks (NN) have emerged as an important tool for classification. It is established that NN are alternative to various conventional classification methods.[8] Neural networks method is used for classification, clustering, featuring mining, prediction and pattern recognition. When output is continuous – prediction and when output is discrete value- classification, and rearrangement of the neurons – detecting clustering. [4]

To implement NN as a classifier for the prediction of learning disability problem, the main concept of Multi-layer perceptron with back propagation is implemented using weka tool. Back propagation is the most widely used learning method. The method of learning involves modifying the weights and biases of the network in order to minimize a cost function.[3] The back propagation algorithm is the workhorse of learning in neural networks**.** Back propagation is an expression for the partial derivative ∂C/∂w of the cost function C with respect to any weight w (or bias b) in the network.[1] It actually gives us detailed insights into how changing the weights and biases changes the overall behaviour of the network. The data has contains 1020 cases having symptoms of learning disability and the 16 attributes are taken from signs and symptoms that are present in the general checklist for the informal assessment of LD. These inputs are given to the system and the output of the system is LD true or LD false. [4]

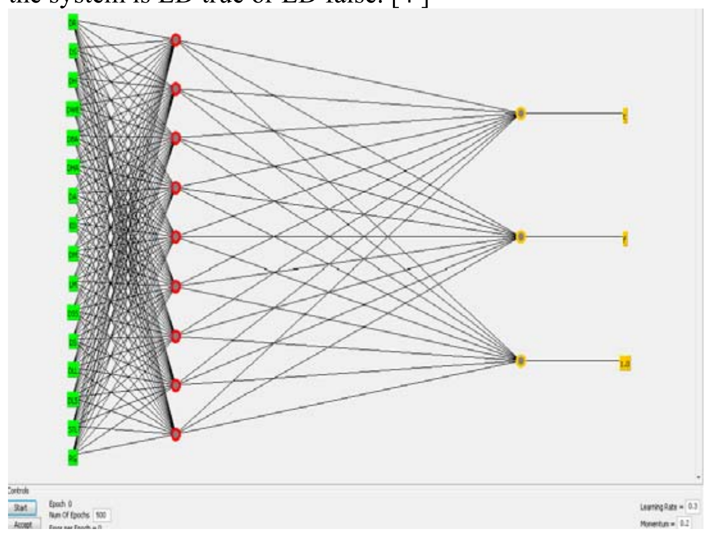

Figure 4 Architecture of MLP network.

The above figure output preformed in weak tool and result analysis are, They are 3 layers in the NN – input are 16 attributes , 9 hidden neurons and 3 output nodes. In the output nodes 2 nodes are true and false. The learning rate Obtained is 0.3, and epoch-500 and momentum 0.2. Error of epoch =0. Activation function-considered for each node in the network is the binary sigmoid function.[4]

Sigmoid : A sigmoid function is a bounded differentiable real function that is defined for all real input values and has a positive derivative at each point and is mathematical function having an "S" shape (**sigmoid curve**).[1] Defined by the formula is.

$$S(t) = \frac{1}{1 + e^{-t}}.$$

Table - 2 Classifier model full training set

IJISET - International Journal of Innovative Science, Engineering & Technology, Vol. 1 Issue 8, October 2014.
www.ijiset.com
ISSN 2348 – 7968

Sl. No	Sigmoid	Threshold
1	Node 0	-0.016241901923750777
2	Node 1	-0.598040902392985
3	Node 2	-1.3241745517865613
4	Node 3	-0.5541893062773827
5	Node 4	0.24625321785824697
6	Node 5	0.21384450430679258
7	Node 6	0.5757657748550636
8	Node 7	-0.479722035810239
9	Node 8	0.44695728552405717
10	Node 9	0.4199671265900112
11	Node 10	0.20015351031343415
12	Node 11	-1.1056611347641336

A decay parameters causes to the learning rate. Here one hidden layer is used. The number of hidden neuron is reduced is the main challenge in the training phase and determines the appropriate number of hidden is the experimentation. The classifier model of full training set and the stratified cross-validation summary obtained while implementing MLP in weka on 1020 data set.

Table- 3 Stratified cross-validation summary obtained

Sl. No	Particulars	Neural Network value
1	Correctly Classified instances	986(96.666%)
2	Incorrectly Classified instances	34.(3.333%)
3	Kappa statistic	0.9329
4	Mean absolute error	0.03164
5	Root mean squared error	0.1056
6	Relative absolute error	6.174%
7	Root relative squared error	29.5363%
8	Total Number of Instances	1020
9	Time taken to build model	29.2 Seconds

The different types of error functions used in this NN are mean squared error, mean absolute error, root mean squared error, relative absolute error and root relative squared error, the values of which are also shown in Table 3.

The major issue studied from this study of prediction of LD in children is failure of the classifier in handling the missing values in the datasets. The missing values contribution may be some times very important and significant. The second issue noticed is that some of the attributes in the check list have less contribution in LD prediction. So they have to reduce the number of attributes for improving the performance of the classifier. Reducing the number of attributes is very effective and that will help to reduce the time taken for constructing the model.

The results obtained shows that 96.67% accuracy with correctly classified instances and 3.33% accuracy in incorrectly classified instances. The findings show that there is no solution in the case of missing values present, also some attributes are unwanted and hence have no contributions in predicting the LD.[4] But from the output of this classification method, it is understand how easily the learning disability can be predicted in the early stages itself.

2.1 Support Vector Machine

Anuradha et al [3] proposed diagnosis of ADHA (Attention Deficit Hyperactivity Disorder) using SVM algorithm. This performed in CLEMENTINE 12.0, Tool.

Support vector machines (SVMs) are a set of related supervised learning methods used for classification and regression. A support vector machine constructs a hyperplane or set of hyperplanes in a high-dimensional space, which can be used for classification, regression or other tasks [9]. To improve on the overall identification accuracy; they also make use of the GA-based, Feature Selection Algorithm. Genetic algorithms are known to give good solution to very complex problems, expect that AI techniques like SVM will certainly play an essential role in future ADHD diagnosis applications.

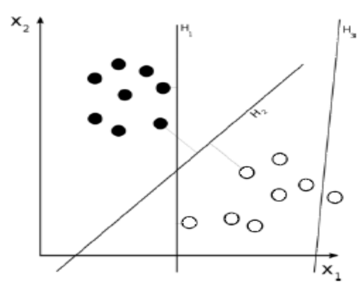

Figure 5 Support vector Machine Representation

Intuitively, a good separation is achieved by the hyperplane that has the largest distance to the nearest training data points of any class (so-called functional margin), since in general the larger the margin the lower the generalization error of the classifier [10].

The training data used is in the form:

D = {(xi,ci) | xi ε R p, ci ε {-1,1}}i=1 to n ...(1)
The SVM is given by :

(w.x1) + b = +1 (2)
(w.x1) + b = -1

=> (w. (x1 – x2)) = 2

=> ((w/||w||).(x1-x2)) = 2/||w|| (3)

Parameter *C* determines the trade off between the model complexity (flatness) and the degree to which deviations larger than ε are tolerated in optimization formulation

for example, if *C* is too large (infinity), then the objective is to minimize the empirical risk only, without regard to model complexity part in the optimization formulation [11]

Parameter ε controls the width of the ε -insensitive zone, used to fit the training data. The values of can ε affect the number of support vectors used to construct the regression function. The bigger ε, the fewer support vectors are selected. On the other hand, bigger ε -values results in more "flat" estimates. Hence, both *C* and ε -values affect model complexity (but in a different way).

ADHD and its diagnosis and treatment have been considered controversial since the 1970s. The controversies have involved clinicians, teachers, policymakers, parents and the media. Opinions regarding ADHD range from not believing it exist at all to believing there are genetic and physiological bases for the condition as well as disagreement about the use of stimulant medications in treatment [12].
To implement this algorithm they used a tool called Clementine 12.0. Clementine is a graphical interface puts the power of data mining in the hands of the user. They describe how use this tool for our benefit i.e. in order to diagnose ADHD [13].

General Algorithm

Procedure to diagnose ADHD in student using SVM

Collect data set = {data set 1, data set 2…..} spilt the data set into T and D.

T: contains the trained data set with diagnosis &

D: contains the data set to be diagnosed

Repeat the next step for data sets T and D

Apply pre-processing to reduce the noisy data

Create a stream in Clementine 12.0 with SVM algorithm

Radial basis function kernel. Apply this stream to the trained data set T. Now, apply the stream to both data set T and D combined.

Output is the diagnosis of ADHD

The general algorithm described and helps us in understanding the steps and analysis of data. The steps taken by us are explained in brief below:

Step 1: (Collection of Data) this step deals with the collection of data and representing the data in the form of an MS-EXCEL SHEET.

Step 2: (Test Data Processing) this step deals with using the SVM tool called CLEMENTINE 12.0 and processing the data through it. Thus, this step deals with the test data.

Step 3: (Diagnosing) in this step, the testing and actual diagnosing takes place. [3]

They introduced the data that needs to be diagnosed and run it through the SVM tool again.

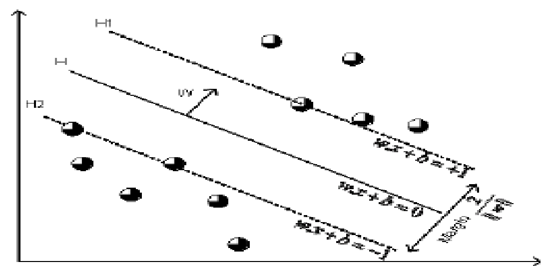

Figure-6 Output given by SVM

To use the model in the Clementine tool they created a stream to work on and create SVM algorithm module for Radial basis function kernel type. After executing this database according to the support vector machine it provides us with its own interpretation of the results. On executing the module the results. Provided by this module are in tabular form (can be graphical too) with a last column added which gives us the algorithm's diagnosis. For the data collected it had an 88.3% of efficiency [14]. This shows us that out of a 100 values for test data they got an accuracy of 88% in diagnosis which is highly encouraging at this stage of research.

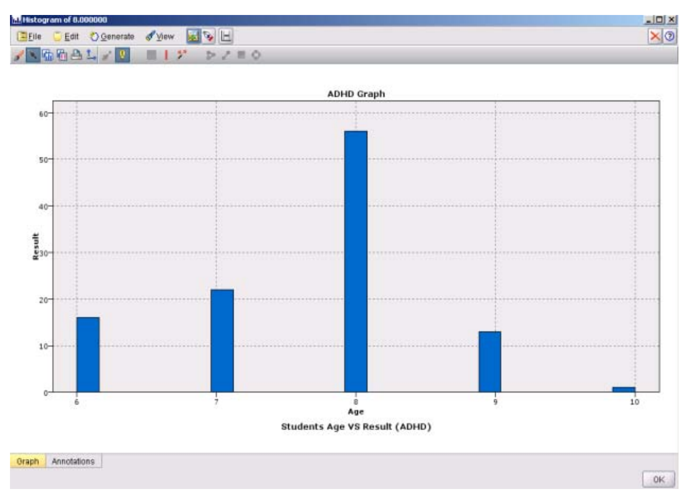

Figure- 7 Graph showing the occurrence of ADHD which age as predicted by SVM algorithm

IJISET - International Journal of Innovative Science, Engineering & Technology, Vol. 1 Issue 8, October 2014.
www.ijiset.com
ISSN 2348 – 7968

The major issue on this study Fig-7 shows the occurrence of ADHD in children between the ages 6 to 11 yrs as evaluated by the SVM algorithm. According to this graph number of reported and diagnosed cases with the child's age 8 yrs are maximum followed by children of age 7. Study about diagnosing ADHD using SVM algorithm has shown a percentage of 88.674% success in diagnosing. There a lot of scope for development in this field of study in order to increase the percentage success achieved by using this technique alone.

2.3 Decision Tree

Kannan Balakrishnan et al [4] proposed Implementation on classification method such as Decision Tree machine. Decision tree builds classification or regression models in the form of a tree structure. It breaks down a data set into smaller and smaller subsets while at the same time an associated decision tree is incrementally developed. The final result is a tree with decision nodes and leaf nodes.[1] A decision node –specifies some test to be carried out on a single attribute-with one branch and sub tree for each possible outcome of the test., where each internal node denotes a test on an attribute, each branch of the tree represents an outcome of the test and each leaf node holds a class label [16]. The topmost node in a tree is the root node class label represents a classification or decision.

The core algorithm for building decision tree called ID3 by J.R. quinlan which employs a top-down greedy search through the space of possible branches with no backtracking. ID3 uses Entropy and information gain to construct a decision tree. [1].

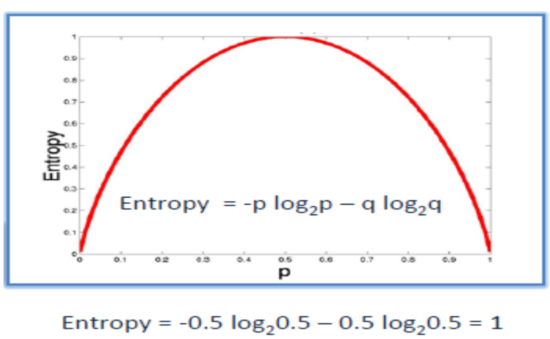

Figure-8 Entropy

The key requirements to do mining with decision tree are: attribute value description, predefined classes, discrete classes and sufficient data.

Data mining techniques are useful for predicting and understanding the frequent signs and symptoms of behavior of LD. If study the attributes of LD, they can easily predict which attribute is from the data sets more related to learning disability. The first task to handle learning disability is to construct a database consisting of the signs, characteristics and level of difficulties faced by those children. Then, Decision or Classification Tree is a tree associated with LD such that each internal node is labeled with attributes DR, DS, DH, DWE, etc.

Each arc is labeled with predicate, which can be applied to the attribute at the parent node. The basic steps in the decision tree are building the tree by using the training data sets and applying the tree to the new data sets.

(i)Attribute List- DR,DS, DH DWE, DBA, DHA, DA, ED, DM, LM, DSS, DNS, DLL, DLS, STL and RG

(ii) Attribute Selection Method by Gain Ratio

The Information Gain Ratio for a test is defined as follows. IGR (Ex, a) = IG / IV, where IG is the Information Gain and IV is the Gain Ratio. Information gain ratio biases the decision tree against considering attributes with a large number of distinct values. So it solves the drawback of information gain.

(iii) Classification

As shown in the stratified cross-validation summary 98.24% cases are correctly classified and 1.76 % cases are incorrectly classified.

This decision tree can be used for making classifications. Here the J48 algorithm is used, which is a greedy approach in which decision trees are constructed in a top-down recursive divide and conquer manner. To illustrate this method, first partition the datasets into two subsets and choose one of the subsets for training and other for testing.[17] Then swap the roles of the subsets so that the previous training set becomes the test set and vice versa.

Table – 4 Stratified cross-validation summary obtained

Sl. No	Particulars	Decision Tree value
1	Correctly Classified instances	1002 Nos.98.2353 %
2	Incorrectly Classified instances	18 Nos., 1.7647%
3	Kappa statistic	0.9627
4	Mean absolute error	0.0263
5	Root mean squared error	0.1317
6	Relative absolute error	5.5664 %
7	Root relative squared error	27.1166 %
8	Total Number of Instances	1020

ISSN 2348 – 7968

9	Time taken to build model	0.22 second

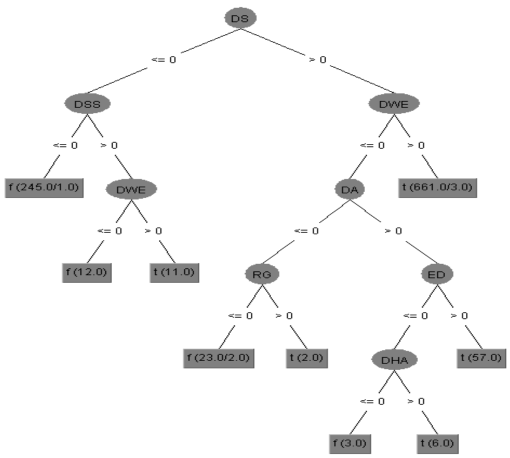

Figure -9 Decision Tree

The rules are so popular because each rule represents an independent knowledge. New rule can added to an existing rule sets without disturbing them, whereas to add to a tree structure may require reshaping the whole tree.

The Rules extracted from the decision tree are given below[18].

R1 : IF DS = NO, DSS = NO THEN LD = NO
R2: IF DS = NO, DSS = YES, DWE = NO THEN LD = NO
R3: IF DS = NO, DSS = YES, DWE = YES THEN LD = YES
R4: IF DS = YES, DWE = NO, DA = NO, RG = NO THEN LD = NO
R5: IF DS = YES, DWE = NO, DA = NO, RG = YES THEN LD = YES
R6: IF DS = YES, DWE = NO, DA = YES, ED = NO, DHA = NO THEN LD = NO
R7: IF DS = YES, DWE = NO, DA = YES, ED = NO, DHA = YES THEN LD = YES
R8: IF DS = YES, DWE = NO, DA = YES, ED = YES THEN LD = YES
R9: IF DS = YES, DWE = YES THEN LD = YES

Confidence and support are properties of rules. These statistical measures can be used to rank the rules and hence the predictions.

Support: The number of records in the training data set that satisfies the rule.

Confidence: The likelihood of the prediction outcome, given that the rule has been satisfied.

Table -5 Confidence of rules table as follows.

Rules	**Confidence**
R1	93%
R2	92%
R3	91%
R4	90%
R5	91%
R6	93%
R7	91%
R8	90%
R9	92%

In decision tree the main objective of attribute evaluation is based on information gain. The wrong predictions obtained from decision tree for all inconsistent data sets can be lead to a limited accuracy of decision tree models. Also the formation of tree and rule generation becomes complex due to the increase of number of attributes. The extracted rules are very effective for the prediction. The wrong predictions obtained from decision tree for all inconsistent data sets can be lead to a limited accuracy of decision tree model. The main drawback noticed from this study is that the failure in handling inconsistent data. Also the formation of tree and rule generation becomes complex due to the increase of number of attributes.

Conclusion:

In this survey paper, the prediction of learning disability in school age children is implemented through various algorithms. The main problem considered, in the work for analysis and solving, is the design of an LD prediction tool based on machine learning techniques. A containing 16 attributes and 1020 real data sets identified signs and symptoms or attributes related to LD are developed and used for the general assessment of learning disability.

A detailed study on the uses of different classification algorithms, viz. Neural Network, Decision Tree, and Support Vector Machine, used for the prediction of learning disabilities in children. The data mining tool Weka, is performed for implementing the Neural network and Decision tree classification algorithms and the SVM is performing in Clementine2.0 tool, in real datasets. The main drawback found in all these classification algorithms is that, there is no proper solution for handling the inconsistent or unwanted data in the data base and also the classifier accuracy is low. Hence, the classification accuracy has to be increased by adopting new methods of implementation by proper data preprocessing.

Reference:

1. Pang-Ning, T., Michael, S., Vipin, K: Introduction to Data Mining,Low price edn. Pearson Education, Inc., London, 2008.
2. Fayyad Usama, Piatetsky Shapiro Gregory, Smyth Padhraic: From Data Mining to Knowledge Discovery in Databases, American Association for Artificial Intelligence, 1996, pp37-54.

3. J Anuradha, Senior Professor - Diagnosis of ADHD using SVM algorithm.
4. Julie m. David Statistical machine learning techniques For the prediction of learning disabilities In school-age children (Final Thesis)
5. Eduardo Ogasawara, Leonardo C. Martinez, Daniel de Oliveira,Geraldo Zimbrao, Gisele L. Pappa, Marta Mattoso: Adaptive Normalization: A Novel Data Normalization Approach for Non- Stationary Time Series, International Joint Conference on Neural Networks, DOI: 10.1109/IJCNN.2012.5596746, IEEE, 2010, pp1-8
6. Julie M. David, Kannan Balakrishnan: Learning Disability Prediction Tool using ANN and ANFIS, International Journal of Transaction on Knowledge and Data Engineering, IEEE. (communicated)
7. Julie M. David, Kannan Balakrishnan: Modified Pre-Processing Methods with Machine Learning Algorithms for LD prediction, International Journal of Human Computer Studies, Elsevier B.V. (communicated).
8. Guoqiang Peter Zhang: Neural Networks for Classification: A Survey, IEEE Transactions on Systems, Man and Cybernetics – Part C: Applications and Reviews, 30(4), Nov. 2000, pp451-462
9. Digital Least Squares Support Vector Machines, David

   Anguita and Andrea Boni. Neural Processing Letters journal Springer publications, volume 18, number 1 august 2003 pages 65-72

10. Davide Anguita, Andrea Boni and Sandro Ridella “Digital Architecture for support vector machines: Theory, Algorithm and FPGA implementation”, Neural Networks, IEEE Transactions on In Neural Networks, IEE Transactions on, Vol. 14, No. 5. (2003), pp. 993-1009

11. Stuart Andrews, Ioannis Tsochantaridis, Thomas Hofmann **“**Support Vector Machine for Multiple-Instance Learning” Eighteenth national conference on Artificial intelligence Edmonton, Alberta, Canada Pages: 943-944,2002

12. Machine Learning Using Support Vector Machines King Saud University The College of Computer & Information Science Computer Science Department (Master) Neural Networks and Machine Learning Applications

13. J.Weston, S.Mukherjee, O.Chapelle, M.Pontil, T.Poggio, V.Vapnik “Feature Selection for SVMs” IEEE paper (2001): 668-674

14. Digital Least Squares Support Vector Machines, David Anguita and Andrea Boni. Neural Processing Letters journal Springer publications, volume 18, number 1 august 2003 pages 65-72

15. Witten Ian, H., Frank Eibe: Data Mining–Practical MachineLearning Tools and Techniques, Morgan Kaufmann - Elsevier
    Publishers, ISBN: 13: 978-81-312-0050-6, 2005
16. Crealock Carol, Kronick Doree: Children and Young People with Specific Learning Disabilities, Guides for Special Education, No. 9, UNESCO, 1993
17. Witten Ian, H., Frank Eibe: Data Mining–Practical MachineLearning Tools and Techniques, Morgan Kaufmann - Elsevier
    Publishers, ISBN: 13: 978-81-312-0050-6, 2005
18. Julie M. David, Kannan Balakrishnan: Prediction of Learning Disabilities in School Age Children using Decision Tree, Proceedings of the International Conference on Recent Trends in Network Communications, Lecture Notes in Communication in Computer and Information Science (CCIS), 90(3), Springer- Verlag Berlin Heidelberg, 2010, pp 533-542
19. http://www.helpguide.org/mental/learning_disabilities.htm
20. Introduction to Clementine, http://www.spss.com or contact