JOURNAL OF INFORMATION, KNOWLEDGE AND RESEARCH IN COMPUTER ENGINEERING

# A FRAMEWORK FOR LEARNING DISABILITY PREDICTION IN SCHOOL CHILDREN USING NAIVE BAYES - NEURAL NETWORK FUSION TECHNIQUE

[1]*K.AMBILI,* [2]*P.AFSAR*

[1, 2] **Department of Computer Science and Engineering,**
**M.E.A. Engineering College, Calicut University,**
**Perinthalmanna, Vengoor, Pattikkad–679325, Malappuram Kerala, India.**

***ambili1992@gmail.com***

***ABSTRACT :*** *A learning disability can cause a child to have some kind of trouble in learning and using certain skills. The skills most commonly affected are reading, writing, listening, speaking, reasoning and doing math. Learning disability prediction is a very complicated task. The purpose of prediction of learning disability during early childhood is very essential since it helps to determine child's strengths and weaknesses and to understand how he or she best learns and where they have difficulty. The information gained from an evaluation is very important for finding out how the parents and the school authorities can provide the best learning environment for child. This paper aims at analyzing various data mining techniques for the prediction of learning disability. The observations show that the fusion technique of naive bayes and neural network is found to be the best among classification and prediction algorithms in the diagnosis of learning disability when compared to other machine learning algorithms.*

***Key Words :*** ***Learning Disability, Naive Bayes Algorithm, Neural Network, Fusion Of Naïve Bayes And Neural Network Classifier****.*

## 1. Introduction

Learning disability is a general term that describes specific kinds of learning problems. Learning disabilities vary from child to child. One child with learning disability may not have the same kind of learning problems as another child with learning disability. A learning disability can cause a child to have trouble in learning and using certain skills. The skills most often affected are: reading, writing, listening, speaking, reasoning and doing math. They usually contain three essential elements: a discrepancy clause, an exclusion clause, and an etiologic clause. The discrepancy clause states there is a significant disparity between aspects of specific functioning and general ability; the exclusion clause states the disparity is not primarily due to intellectual, physical, emotional, or environmental problems, and the etiologic clause speaks to causation involving genetic, biochemical, or neurological factors. The most frequent clause used in determining whether a child has a learning disability is the difference between areas of functioning.

Over half of all children who receive special education have a learning disability . There is no one sign that shows a child has a learning disability . Experts look for a noticeable difference between how well a child does in school and how well he or she could do, given his or her intelligence or ability. There are also certain clues, most relate to elementary school tasks, because learning disabilities tend to be identified in elementary school, which may mean a child has a learning disability . A child probably won't show all of these signs, or even most of them. However, if a child shows a number of these problems, then parents and the teacher should consider the possibility that the child has a learning disability. If a child has unexpected problems in learning to read, write, listen, speak, or do math, then teachers and parents may want to investigate more. The same is true, if the child is struggling to do any one of these skills. The child may need to be evaluated to see if he or she has a learning disability.

When a learning disability is suspected based on parent and/or teacher observations, a formal evaluation of the child is necessary[14]. A parent can request this evaluation, or the school might advise it. Parental consent is needed before a child can be tested . Many types of assessment tests are available. Child's age and the type of problem determines the tests that child needs. Just as there are many different types of learning disabilities, there are a variety of tests that may be done to pinpoint the problem. A complete evaluation often begins with a physical examination and testing to rule out any visual or hearing impairment.

However there is no "cure" for learning disabilities. They are life-long. Children with learning disability can be high achievers and can be taught ways to get around the learning disability. With the right help, children with learning disability can learn successfully and perform better. Hence the main aim of the research is to identify school children with

learning diasbilities at a very young age itself so that they can be given proper assistance.

A major idea of machine learning research is to automatically learn to recognize complex patterns and make intelligent decisions based on data. In recent years the sizes of databases have increased rapidly. This has lead to a growing interest in the development of tools capable in the automatic extraction of knowledge from data. Data mining is a computational method of processing data which is successfully applied in many areas that aim to obtain useful knowledge from the data[5,6]. Data mining techniques are used to build a model according to which the unknown data will try to identify the new information[7]. Regardless of origin, all data mining techniques show one common feature: automated discovery of new relationships and dependencies of attributes in the observed data[1,2]. Data mining consists of several algorithms for clustering, classification, prediction and associations, which try to fit a model closest to the characteristics of data under examination[3].

Therefore, the present study aims to develop an expert system which can evaluate important symptoms which help in the diagnosis of learning disability accurately. In particular the following is the contribution of this work:

• Comparing classification algorithms for learning disability.

• Identifying new attributes for predicting learning disability.

• Improving the accuracy of the algorithm which has shown best results with our dataset.

This paper is organized as follows. Section 2 presents some related work and recent studies on learning disabilities using data mining techniques. Section 3 summarizes data mining algorithms used in this work. Section 4 gives a brief overview of the available data and the transformations carried out to clean and put the data in the proper format for analysis. Section 5 gives the description of the classification algorithms which have shown best accuracy with our dataset. Section 6 presents the obtained results and Section 7 concludes with some remarks about the described work and guidelines for future work.

## 2. RELATED WORKS

As many as 1 out of every 10 children has a learning disability. Almost 3 million children (ages 6 to 16) have some form of a learning disability and receive special education in school . In fact, over half of all children who receive special education have a learning disability. Unfortunately there are not many works done for diagnosing learning diasbility using data mining techniques.Some of the data mining technique used for predicting learning disability used machine learning algorithms such as Rough set apprach and Decision tree algorithm.

The Rough set approach seems to be of fundamental importance to artificial intelligence [8]. Rough set theory (RST) has been successfully applied in many real life problems such as medicine, pharmacology, engineering, banking, finance, market analysis, environment management and others. The rough set approach of data analysis has much important advantage.

During the late 1970s and early 1980s, J. Ross Quinlan, a researcher in machine learning developed a decision tree algorithms known as ID3 [9]. This work expanded on earlier work on concept learning system. Decision tree method is widely used in data mining and decision support system. Decision tree is fast and easy to use for rule generation and classification problems. It is an excellent tool for decision representations. The accuracy of a classifier refers to the ability of a given classifier to correctly predict the class label of new or previously unseen data.

For prediction of learning disability, decision trees are probably the most frequently used tools for rule extraction from data,[10,11] whereas the rough sets based methods seems to be their newer alternative [12,13]. In both cases, the algorithms are simple and easy to interpret by users. The practical aspects of application of those tools are different. The computational times of decision trees are generally short and the interpretation of rules obtained from decision trees can be facilitated by the graphical representation of the trees. RST may require long computational time and may lead to much large number of rules compared to DT. The rules extraction algorithm is very important, particularly in construction of data mining system. Therefore, we have to go for some other machine learning algorithms.

## 3. AVAILABLE DATA

The data has been collected from a special school in Kerala. The collected data comprised of symptoms of children who were suffering from learning disabilities. This collection was made available in a csv file.

**List of Attributes:**

The data included various symptoms of learning disability as attributes. They are :-

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
6. DHA - Difficulty with Higher Arithmetic
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

This data was used as the training set for various algorithms. The testing data was collected through the questionnaire of 30 school children.

## 4. EXPERIMENTAL WORKS

Classification is a supervised learning technique of data mining which assigns items in a collection to a targetclass. Each classification algorithm has a different approach for determining the relationship between the attributes in order to predict the output. The data is partitioned into two sets: one for training the model and second one for testing the model. In our model the dataset used for training consists of 1124 instances and 16 attributes. The performances of various classifiers are observed and the accuracy is assessed.

## 5. ALGORITHMS

**Naive Bayes Algorithm**

Naïve Bayes is one of the most efficient classification algorithm. . It is a simple probabilistic classifier based on applying Bayes theorem. It can be used for solving diagnostic and predictive problems.

The naive Bayes classier greatly simplify learning by assuming that features are independent given class Naive Bayes model records how often a target field value appears together with a value of an input field. It considers each of the symptoms to contribute independently to the probability that the child has learning disability. It estimates the probability of observing a certain value in a given class by the ratio of its frequency in the class of interest over the prior frequency of that class.

**P (x1, x2, x3....xd|C j) = P (xi|C j).....(1)**

**P (c|X ) = P (x1|c) * P (x2|c) * P(x3|c)..... *P (xn|c) *P (c) .....(2)**

For example : If a child has symptom x1(DR), x2 (DS), x3(DH), x4(DWE) then the probability of child having learning disability or not can be calculated through the following process :-

**Step 1**: probability of child having learning disability can be calculated.

**P (x1|C 1)** = number of children having DR and have learning disability / number of children having learning disability.

**P(C1**) = number of children having learning disability / total number of children.

**P (xn|C1) = P (x1|C1) * P (x2|C1) * P (x3|C1) * P (x4|C1) * P (C1)**

**Step 2**: Probability of user not having learning disability can be calculated.

**P (xn|C2) = P (x1|C2) *P (x2|C2) * P (x3|C2) * P (x4|C2) * P (C2)**

**Step 3:** The probability of user having or not having learning disability has been compared.

If **P (xn|C1)** is greater then that child is having learning disabiity else vice-versa.

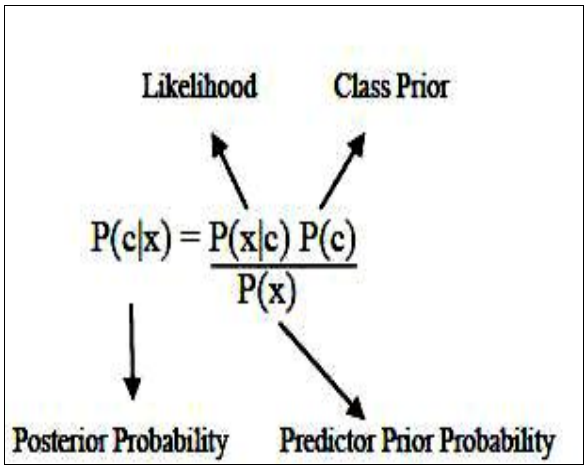


Figure 1 : Naive Bayes Formula

**Neural Network Classifier**

Artificial neural networks are based on the functioning of human brain. It is used when the relationship is complex or unknown between the underlaying data. It consists of three layers: input layer, hidden layer and output layer. Input layer consists of neurons which are independent variables supplied to find correlation between input and output. There are two main functions which play an important role in neural networks: input function and activation function. Input to a node is the weighted sum of outputs of neurons connected to it.

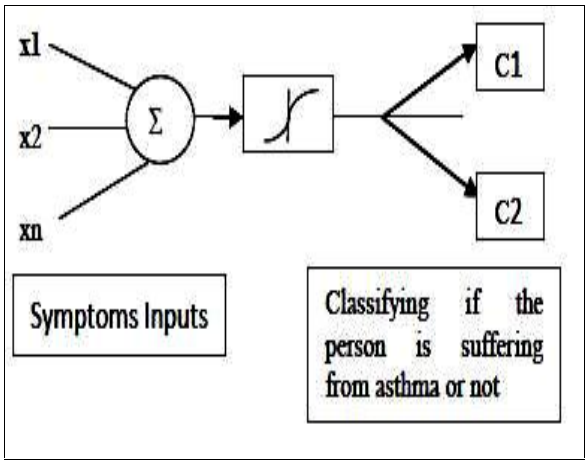


Figure 2 : Neural Network

Neti = wijxj ; output of jth unit. Activation value of the unit, is g(wjixi) , here g(.) is activation function and xi is output of ith unit connected to unit j.

**g(netInput) = 1/(1+(e^netInput)) .....(3).**

The task of activation function is to map output between the interval [J1, J2] . In proposed work, we have used sigmoid activation function as it maps the output into the interval [0, 1], that are nonlinear and continuously differentiable which are desirable for network learning. The weight vector for the ith processing unit at time instant (t+1) in terms of the weight vector at time instant (t) as follows:

**wi (t +1) = wi (t) + wi (t) .....(4)**

In (**4**) wi (t) is the change in the weight vector. The network adapts as follows: change the weight by an amount proportional to the difference between the desired output and the actual output. As an equation:

**Wi = * (D-Y).Ii .....(5)**

We obtain learning rate from (**5**). It influences the speed and quality of learning where,D is the desired output, Y is the actual output, and Ii is the ith input. This is called the Perceptron Learning Rule. At the forecasting stage, one would apply the proper input matrix to the trained network and obtain the learning disability forecast .

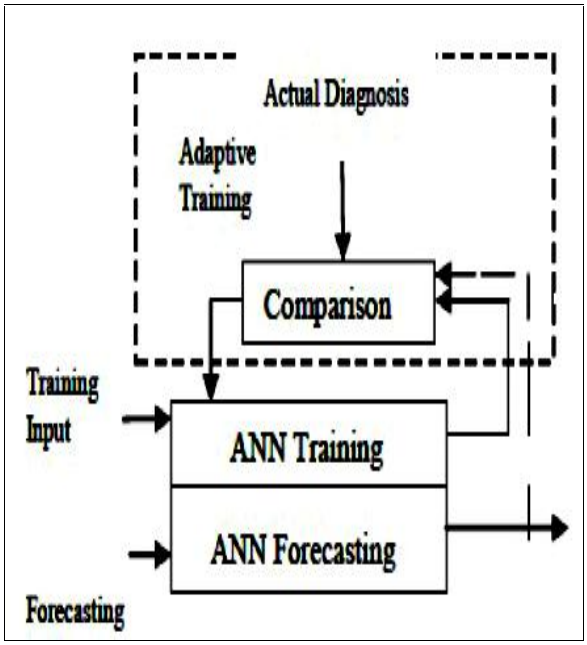


Figure 3 : Learning disability diagnosis through Back propagation

## 6. PROPOSED METHOD – FUSION OF NAIVE BAYES AND NEURAL NETWORK CLASSIFIER

The naive bayes shows more accuracy than backpropagation neural network. The outputs of both the algorithms are taken and are given random weights. The combination in which the best accuracy is achieved, it is calculated.

In this model Pk(i) denotes the blocking probability that K wavelengths are used on the $i^{th}$ link of the path.

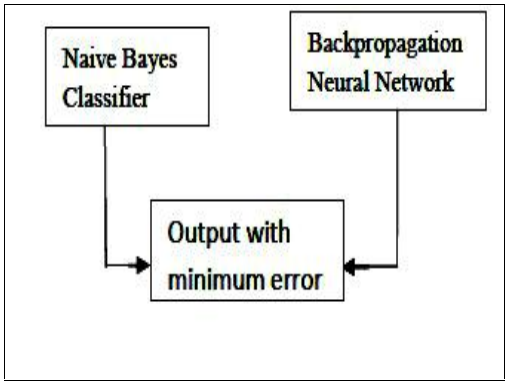


Figure 4 : Fusion of Naive Bayes and Neural Network

## 7. RESULTS

The experiment makes a comparative study on the performances of machine learning algorithms for predicting learning disability. They are evaluated on the basis of three criteria :-

1.Prediction Accuracy
2. Learning Time and
3. Error Rate

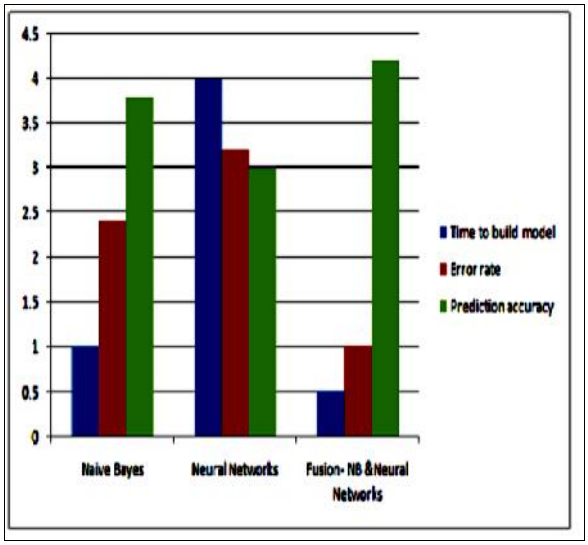


Figure 5 :- Comparison of the three algorithms based on Learning time, Error rate and Prediction accuracy

From the results, it was able to understand that our proposed approach Naive Bayes - Neural Network fusion technique provides more number of correctly classified instances than the other two algorithms. Regarding the Learning time of Algorithms, it was able to understand that Artificial Neural Network consumes more time to build the model. Out of these three algorithms, our proposed method has high prediction accuracy than other algorithms.

## 8. CONCLUSION AND FUTURE WORK

This paper highlights an improved machine learning approach for learning disability prediction in school children using a fusion of the conventional Naive Bayes and Neural Network Classifier**.** While there is still scope for further improvement in the classification time of classical Naive Bayes technique, we believe that we have been successful in our goal of proposing a technique that retains the simplicity of both Naive Bayes and Neural Networks and shows better accuracy in predicting learning disability in school children. This study will be helpful for the parents, teachers and school authorities in diagnosing the child's problem at an early stage. Hence these results will be helpful and beneficial for the educational as well as medical communities.

## 9.REFERENCES:

1. Sellappan Palaniappan and Rafiah Awang, "Intelligent Heart Disease Prediction System Using Data Mining Techniques," IJCSNS International Journal of Computer Science and Network Security, Vol.8, No. 8, August 2008.

2. E. A.Q. Ansari, Neeraj Kumar Gupta, "Automatic diagnosis of asthma using neurofuzzy system," Fourth International Conference on Computational Intelligence and Communication Networks, Vol.7, April 2012.

3. Julie M. David, Kannan Balakrishnan, "Prediction of Frequent Signs of Learning Disabilities in School Age Children using Association Rules," In Proceedings of the International Conference on Advanced Computing, Vol 13, April 2009.

4. Mansi Gera and Shivani Goel, "A Model for Predicting the Eligibility for Placement of Students Using Data Mining Technique," International Conference on Computing, Communication and Automation, pp. 18-23, Vol. 04, Issue 01, January 2015.

5. Muhamad Hariz and Wahidah Husain "A Framework for Childhood Obesity Classifications and Predictions using NBtree," International Conference on IT in Asia, No.8, November 2011.

6. B. Novak, M. Bigek, "Application of artificial neural networks for childhood obesity prediction", IEEE, Slovenia, 1995, pp. 377-380.

7. Cios, K.J., Pedrycz W., Swiniarski, R.W. and Kurgan, L.A., "Data Mining: A Knowledge Discovery Approach", Springer, New York.

8. Grzymala-Busse JW, " Knowledge Acquisition under Uncertainty-A Rough Set Approach," Journal of Intelligent & Robotic Systems, Vol 1, 3-16, 1988.

9. Quinlan J.R., 1986. Induction on decision trees, Machine learning, 1(1):81-106.

10. Chen R.S, Wu R.C., Chang C.C, "Using data mining technology to design an intelligent CIM system for IC manufacturing," in proceedings of sixth international conference on software engineering, artificial intelligence, network, parallel distribution, computation self assembly wireless network, 70-75, 2005

11. Stuart R., Peter N., "Artificial Intelligence – A Modern approach," Pearson Prentice Hall, 2009.

12. Kusiak A, Kurasek C, "Data mining of printed circuit board defects", IEEE Trans, Rob Autom, Vol 17, No.2, 74-384, March 2006

13. Tseng T.L., Jothishankar M.C., Wu T., Xing G., Jiang F, "Applying data mining approaches for defect diagnosis in manufacturing", World Academy of Science, Engineering and Technology, 61, 2010

14. Blackwell Synergy, " Learning Disabilities Research Practices", Volume 22, 2006