International Journal of Computer Applications (0975 – 8887)
Volume 9– No.11, November 2010

Machine Learning Approach for Prediction of Learning Disabilities in School-Age Children

Julie M. David
Dept. of Computer Applications
MES College, Marampilly, Aluva
Cochin- 683 107, India

Kannan Balakrishnan
Dept. of Computer Applications
Cochin University of Science & Technology
Cochin - 682 022, India

ABSTRACT
This paper highlights the two machine learning approaches, viz. Rough Sets and Decision Trees (DT), for the prediction of Learning Disabilities (LD) in school-age children, with an emphasis on applications of data mining. Learning disability prediction is a very complicated task. By using these two approaches, we can easily and accurately predict LD in any child and also we can determine the best classification method. In this study, in rough sets the attribute reduction and classification are performed using Johnson's reduction algorithm and Naive Bayes algorithm respectively for rule mining and in construction of decision trees, J48 algorithm is used. From this study, it is concluded that, the performance of decision trees are considerably poorer in several important aspects compared to rough sets. It is found that, for selection of attributes, rough sets is very useful especially in the case of inconsistent data and it also gives the information about the attribute correlation which is very important in the case of learning disability.

Keywords
Decision Tree, Learning Disability, Rough Sets, Rule Mining and Support and Confidence

1. INTRODUCTION
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. The term Data Mining or Knowledge Discovery in databases has been adopted for a field of research dealing with the automatic discovery of implicit information or knowledge within databases [24]. Knowledge Discovery in Databases (KDD) is the process of identifying useful information in data [20]. A widely accepted formal definition of data mining is given subsequently. According to this definition, data mining is the non-trivial extraction of implicit previously unknown and potentially useful information about data [7]. Conventionally, the information mined is denoted as a model of the semantic structure of the datasets. The model might be utilized for prediction and categorization of new data [6]. Diverse fields such as marketing, customer relationship management, engineering, medicine, crime analysis, expert prediction, web mining and mobile computing besides others utilize data mining [12]. A majority of areas related to medical services such as prediction of effectiveness of surgical procedures, medical tests, predication and the discovery of relationship among clinical and diagnosis data also make use of data mining methodologies [3].

This paper presents study of multi-classification methods, viz. rough sets and decision trees and shows how these ideas may be utilized for data mining. The rough set approach seems to be of fundamental importance to artificial intelligence [9]. Rough set theory (RST) has been successfully applied in many real life problems in 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 [22]. 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.

In principle, for extraction of logic rules from data, any classification system or model can be used. Typical learning algorithms include direct rule induction, decision trees, Bayesian classifier, etc. The details on these methods can be found in [18]. Many classifiers have also been successfully utilized for logic rule extraction. This approach facilitates processing continuous value variables handling uncertainties appearing in data.

For prediction of LD, decision trees are probably the most frequently used tools for rule extraction from data,[25,4] whereas the rough sets based methods seems to be their newer alternative [18,27]. In both cases, the algorithms are simple and easy to interpret by users. The practical aspects of application of those tools are different. The computation 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 times 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, there are very little comparative studies are available.

The purpose of the present paper is to show the important differences in performance of these two data mining methods for the prediction of LD. The remaining paper is organized as follows. Section 2 describes about LD. We then provide a brief review on the two machine learning approaches and their results in section 3. Section 4 presents the comparison of results followed by the rule extraction for the prediction of LD in Section 5. Then in Section 6, the result analysis is explained. Finally, Section 7 dealt with conclusion and future research work.

2. LEARNING DISABILITY

Learning disability is a general term that describes specific kinds of learning problems. Learning disabilities are formally defined in many ways in many countries. However, they usually contain three essential elements: a discrepancy clause, an exclusion clause, and an etiologic clause [15]. 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 [17]. When a person shows a great disparity between those areas of functioning in which she or he does well and those in which considerable difficulty is experienced, this child is described as having a learning disability [14]. 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 [15]. Learning disabilities vary from child to child. One child with LD may not have the same kind of learning problems as another child with LD. There is no "cure" for learning disabilities [23]. They are life-long. However, children with LD can be high achievers and can be taught ways to get around the learning disability. With the right help, children with LD can and do learn successfully [14].

As many as 1 out of every 10 children in the United States has a learning disability. Almost 3 million children (ages 6 through 21) have some form of a learning disability and receive special education in school [3]. In fact, over half of all children who receive special education have a learning disability [5]. There is no *one sign* that shows a child has a learning disability [17]. 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 [15]. 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 [16].

When a LD is suspected based on parent and/or teacher observations, a formal evaluation of the child is necessary. A parent can request this evaluation, or the school might advise it. Parental consent is needed before a child can be tested [14]. 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 LDs, 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 [5]. Many other professionals can be involved in the testing process. The purpose of any evaluation for LDs is to determine child's strengths and weaknesses and to understand how he or she best learns and where they have difficulty [15]. The information gained from an evaluation is crucial for finding out how the parents and the school authorities can provide the best possible learning environment for child [14]. Here we are using a checklist for assessing the LD. It contains 16 most frequent signs and symptoms of LD. These symptoms, which are the attributes in this study, are listed in table 1 below.

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

3. MACHINE LEARNING APPROACHES - ROUGH SETS AND DECISION TREES - FOR PREDICTION OF LD

Rough set theory is a new intelligent mathematical tool introduced by Z. Pawlak in 1982[21,10,8]. Rough set theory represents an objective approach to imperfections in data. As per this theory, there is no need for any additional information about data and hence no feedback from additional expert is necessary. All computations are performed directly on data sets [20]. A rough set is an approximation tool that works well when in environments heavy with inconsistency and ambiguity in data or involving missing data [2]. Along the years, rough set theory has earned a well-deserved reputation as a sound methodology for dealing with imperfect knowledge in a simple though mathematically sound way [1].

The decision is a flow chart like structure, 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 [11,16,17]. The topmost node in a tree is the root node. Decision trees are powerful and popular tool for classification and prediction. It is a classifier in the form of a tree structure where each node is either a leaf node-indicates the value of the target attribute of examples or 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 [29]. Classifiers do not require any domain knowledge or parameter setting and therefore is appropriate for exploratory knowledge discovery. Decision tree can handle high dimensional data [16]. The learning and classification step of decision tree are simple and fast. A decision tree can be used to classify an example by starting at the root of the tree and moving through it until a leaf node, which provides the classification of the instance [26].

A tree building process starts by selecting an attribute to place at the root node and at each succeeding level the subset generated by proceeding levels are further partitioned until it reaches a relatively homogeneous terminal node or leaf node. The condition attribute, that induces most amount of entropy reduction and information gain are placed closer to the root node.

When we study decision tree model, we can see that some of the combinations appearing in the data set may be absent in the tree. The lack of combinations of input variables in decision trees, which are present in training data, may result in the rule system in which some important rules are missing. Another thing, sometimes the decision trees can give wrong predictions when inconsistent data are present. In the case of LD, wrong prediction result will make a large problem. So we will consider the solution for recovering that problem and use the simplicity of decision tree structure.

3.1 Results from rough sets

The rough set application development consist four steps. The first step is the development of decision table. In our study, decision table include 513 objects or cases of LD. For each case, 16 attributes are registered. The second step is the approximation of decision space. Here the approximation of object's classification is evaluated. This includes construction of approximation of each decision class with respect to all the condition attributes. The quality of approximation, accuracy and entropy measures are equal to 1. The third one is the reduction of attributes. The extraction of reduct from data involves construction of minimal subset of attributes ensuring the same quality of sorting as that of all attributes. The last step is the rule extraction. It is relatively a straightforward procedure. Reducts are used to generate decision rules from a decision table. The objective is to generate basic minimal covering rules or minimal number of possible shortest rules covering all the cases. The LEM1 algorithm is used to derive minimal sets of rules covering all the objects from learning sets. The algorithm generates six rules that predict the learning disability. Rough Set based algorithm will find a rule with reduced number of conditions, so that they include only those combinations of input values which appear in the data.

The reduct (core attributes) and classification results of the study, on the 513 real data sets with 16 attributes are obtained from Rosetta, the rough set tool kit for analysis of data is shown in Table 2 and Table 3 below respectively. In Rosetta tool, Johnson's reduction algorithm is used for obtaining the reduct results and Naive Bayes Batch classifier is used for obtaining the classification results.

Table 2 . Core attributes

No name

	Reduct	Support	Length
1	{DH, DBA, LM, DSS, STL}	97	5

Table 3. Classification results

No name

		Predicted		
		t	f	
Actual	t	287	30	0.905363
	f	4	192	0.979592
		0.986254	0.864865	0.933723
ROC	Class	t		
	Area	0.985048		
	Std. error	0.004927		
	Thr. (0, 1)	0.136		
	Thr. acc.	0.136		

Our study under RST consists of two parts. The first part is the determination of core attributes (reducts) using Johnson's reduction algorithm and second part is the classification of data into predict LD as *yes* or *no* using Naive Bayes batch classifier algorithm using the Rosetta tool. The major findings from this study are the determination of core attributes of LD, the accuracy of rough set classification and the importance of rule mining for LD prediction in children.

As a pre-processing before data mining, a subset of original data, which is sufficient to represent the whole data set, is generated from the initial detailed data contained in the information system. This subset contains only minimum number of independent attributes for prediction of LD. This attribute is used to study about the original large data set. It is common to divide the database into two parts for creating training set and test set. One of these parts, for instance 10% of the data, is used as training set and examined by the data mining system. The rest of the original database is used as test set for checking whether the knowledge acquired from the training set is general or not. By examining the 513 data in the database, the system tries to create general rules and descriptions of the patterns and relations in database to gain knowledge, which is valid not only in the specific database considered but also for other similar data.

The knowledge is tested against the test set. It is then clearly seen that the patterns found in the training set are valid also for other data. Therefore, if the knowledge gained from the training set is the general knowledge, it is correct for most parts of the test set as well. The learning disability detection process can be considered as a decision making process. The rules generated by considering the original data set give a strong platform for making decisions. We are interested in applying these rules for making decisions.

From this study, we found that RST is more suitable and accurate in selecting attributes. For construction of decision tree, selection of attribute is very important. The rough set theory has been used for selecting attributes, consequently a reduct of attribute will be found which is regarded as a best reduction of attribute and the attribute within this reduct are used for depict the data. The goal is to reduce the volume of data. Here about 16 attributes of LD are used and using reduction algorithm we reduced that into 5. It is a very important thing in the case of LD prediction.

3.2 Results from decision tree

Decision tree induction is one of the simplest, and yet most successful forms of learning algorithm. It serves as a good introduction to the area of inductive learning, and easy to implement [28]. A decision tree takes as input an object or situation described by a set of attributes and returns a decision. A divide and conquer approach to the problem of learning from a set of independent instances leads naturally to a style of representation called decision tree [29]. The basic idea behind the decision tree learning algorithm is to test the most important attribute first; by most important we mean the one that makes the most difference to the classification of an example. That way, we get to the correct classification with a small number of tests, meaning that all paths in the tree will be short and the tree as a whole will be small [28].

We used J48 algorithm in weka, a machine learning workbench, which include a framework in the form of Java class library [13]. Initially we evaluate the worth of an attribute by measuring the information gain ratio with respect to the class. Attributes are then ranked by their individual evaluations by using in conjunction with gain ratio, entropy, etc. In this study, we are using the J48 algorithm for constructing the tree and that model correctly classified 97.47% instances from the data sets using weka. The accuracy of decision trees is given in Table 4 below. The decision tree formed based on the 513 data set is shown in figure 1 below.

Table 4. Accuracy of decision tree

TP Rate	FP Rate	Pre-cision	Recall	F Mea-sure	ROC Area	Class
0.984	0.030	0.981	0.984	0.983	0.968	T
0.979	0.022	0.964	0.979	0.972	0.969	F
Correctly Classified Instances					500 Nos.	97.47%
Incorrectly Classified Instances					13 Nos.	2.53%
Time taken to build a model						0.08Sec

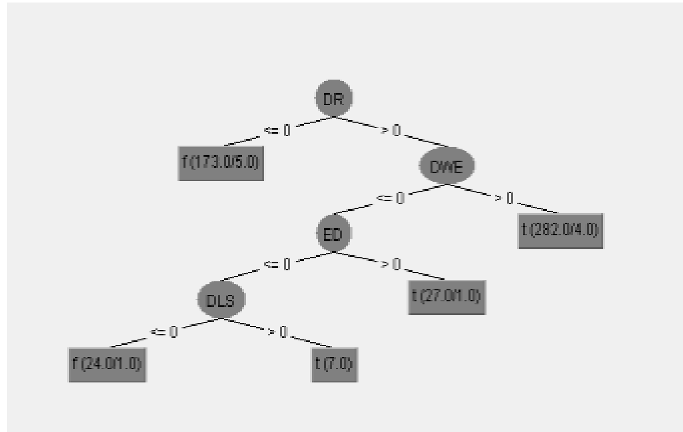

Figure 1. Decision tree

4. COMPARISON OF RESULTS

The result of this study is compared with the results of our other similar studies conducted based on Naive Bayes, Support Vector Machines (SVM) and Muliti Layer Perceptron (MLP) classifiers. The comparison of results is shown in Table 5 below. From this comparison, we can see that rough set is better in terms of classification and accuracy. However, the time taken for building the model is slightly higher in rough sets compared to other classifiers except SVM and MLP, where it is too high.

Table 5. Comparison of classification results

Parameters	Classifiers				
	Rough Sets (Rosetta)	J 48	Naive Bayes	SVM	MLP
Correctly Classified Instances(Nos.)	506	500	426	502	502
Incorrectly Classified Instances (Nos.)	7	13	87	11	11
Time taken to build a model (in seconds)	0.13	0.08	0.06	3.92	18.03
ROC area	0.985	0.968	0.977	0.980	0.980

5. RULE EXTRACTION FOR PREDICTION OF LD

The general requirement for knowledge extraction could be useful in prediction of LD. The rulers should be reliable, which means that, there is a real chance that the application of a rule will bring the predicted result. This result can be expressed by rules quality parameters or the accuracy of the rules of which the most important are support and confidence [19]. Support is defined as rate of observation basis of the rule. Confidence is defined as a ratio of number of records with given combination of input values and the given output value to the number of records which have that combination of input values only. It estimates the probability that application of input values appearing in the rule will give the result expressed by the decision class. The next requirement is that the rules should not be unnecessarily demanding. The requirement and knowledge extraction tools suitable for predicting the LD should specify the available data. If we generating the rules, the rule system should consider (i) the rules should make use of all information in data, (ii) the rule is not redundant and (iii) the rule should be reliable.

From the rough set classification method, under this study, we have mined the following six rules;

R1: (DH,Y) (DBA ,Y) (LM ,Y) (DSS, Y) (STL ,Y)=> (LD,Y) (1)

R2: (DH, N) (DBA, N) => (LD, N) (2)

R3: (DH, Y) (DBA, Y) => (LD, Y) (3)

R4: (DH, N) (DBA, Y) => (LD, Y) (4)

R5: (DBA, Y) => (LD, Y) (5)

R6: (DBA, N) => (LD, N) (6)

From the decision tree classification method, under this study, we have extracted the following seven rules;

R1:(DR=N, DA=N) => (LD, N) (1)

R2:(DR=N, DA=Y, DH=Y)=>(LD,Y) (2)

R3:(DR=N, DA=Y, DHA=N) => (LD, N) (3)

R4:(DR=Y, DBA=N, DLS=N, DSS=N)=>(LD, N) (4)

R5: (DR=Y, DBA=N, DLS=Y) => (LD, Y) (5)

R6: (DR=Y, DBA=N, DLS=N, DS=Y)=>(LD,Y) (6)

R7: (DR=Y, DBA=Y) => (LD, Y) (7)

6. RESULT ANALYSIS

We can see that, both methods provide algorithm for evaluating conditioning attribute, but their inherent significance is entirely different. In decision tree, the main objective of attribute evaluation is based on information gain, while in the concept reduct in rough set, it is based on elimination of redundant attribute in a decision table. The focus is to identify the minimal set of attribute that preserves the indiscernibility relation.

The wrong predictions obtained from decision trees for all consistent and inconsistent data sets can be lead to a limited accuracy of decision tree models. Decision trees have pointed at the decision classes, which are not predominant for the given combination of input values like inconsistent data. The result of this study indicates that the rules system represented by the decision trees may be significantly incorrect for inconsistent data as well as for consistent data with large number of variables. The confidence level of the rules of decision trees shows lower accuracy compared to rough set theory. The quality parameters of rules obtained for RST and DTs with consistent and inconsistent data are represented in Figure 2 below.

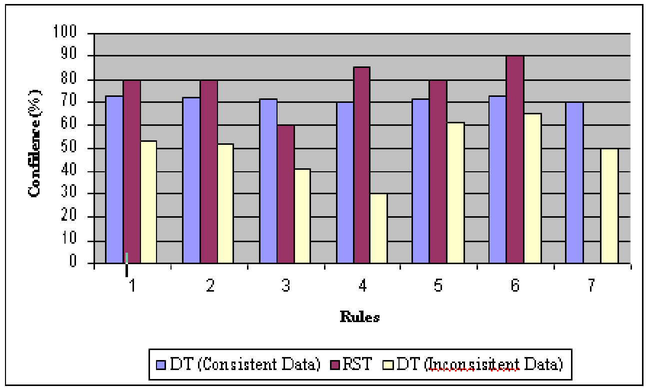

Figure 2. Quality Parameters of Rules

The confidence of the rules, based on RST, shows a higher performance, as shown in Table 5 below, as compared to DT with consistent data.

Table 5. Confidence of rules (RST)

Rules	Confidence
R1	80%
R2	80%
R3	60%
R4	85%
R5	80%
R6	90%

In contrast with decision trees, rough set theory is able to produce different rules, which provide good confidence and support. Rules obtained from rough set theory may not include redundant data. The inconsistent data may lead to false attribute selection in the case of decision tree. In this paper, we are using the information gain as the attribute selection method in decision tree. But the inconsistency of the data leads to the false determination of attribute. In the case of attribute selection rough set is more suitable. The rules obtained from decision trees and rough set theory can offer prediction of LD for combinations of input values absent in data. Here, the input values considered as the symptoms of LD. So the decision trees and rough set theory consider the inconsistent data in different ways. In the case of decision trees, such values may lead to prediction, which is a good reflection of the general dependencies in training data, and the prediction, which is far from the expectations and impossibility of the prediction. The confidence of rules obtained for consistent data in DT is shown in Table 6 below. If the same rules applied on the inconsistent data, the confidence of the rules is reducing to a poor level as shown in Table 7 below.

Table 6. Confidence of rules (DT with consistent data)

Rules	Confidence
R1	73%
R2	72%
R3	71%
R4	70%
R5	71%
R6	73%
R7	70%

Table 7. Confidence of rules (DT with inconsistent data)

Rules	Confidence
R1	53%
R2	52%
R3	41%
R4	30%
R5	61%
R6	65%
R7	50%

From the comparison of results, we have noticed that RST with Naïve Bayes algorithm has a number of advantages over DT for solving the similar nature of problems. For large data sets, there may be chances of some incomplete data or attributes. In data mining concept, it is difficult to mine rules from these incomplete data sets. But in RST, the rules formulated will never influenced by any such incomplete datasets or attributes. Hence, LD can accurately be predicted by using RST method. The other advantage of rough set concept is that it may act as a knowledge discovery tool in uncovering rules for the diagnosis of LD affected children. The importance of RST in this study is that, using a single attribute, we can predict whether a child has LD or not. The sixth rule in RST, which shows 90% confidence, contains only one attribute, which is the most important symptom of LD. If RST is comparing with decision trees, the data or the output of decision tree is very userfriendly. Another thing is that, the output of decision tree is categorical. In comparing the classification results, rough set approach is found better than decision tree approach in correctly classified instances and ROC curve.

7. CONCLUSIONS AND FUTURE RESEARCH

This paper highlights the two machine learning approaches, viz. Rough Sets and Decision Trees, to predict the learning disabilities in school age children. In rough sets the attribute reduction and classification are performed using Johnson's reduction algorithm and Naive Bayes algorithm respectively for rule mining and J48 algorithm is used in construction of decision trees. The extracted rules in both the methods are very effective for the prediction. The wrong predictions obtained from decision trees for all inconsistent data sets can be lead to a limited accuracy of decision tree models. Decision trees have pointed at the decision classes, which are not predominant for the given combination of input values like inconsistent data. The result of this study indicates that, the rules system represented by the decision trees may be significantly incorrect for inconsistent data with large number of variables. The computation times of decision tree are generally short and the interpretation of rules obtained from decision tree can be facilitated by the graphical representation of the trees. The rough set theory may require long computational times and may lead to much large number of rules compared to decision tree. It is found that, for selection of attributes, rough sets is very useful especially in the case of inconsistent data and it also gives the information about the attribute correlation which is very important in the case of learning disability. The results obtained from this study is compared with that of other classifiers such as Naive Bayes, SVM and MLP and it is found that rough set is better in terms of classification and accuracy.

Obviously, as the school class strength is 40 or so, the manpower and time needed for the assessment of LD in children is very high and may not be accurate. But by using the rules developed by us using these approaches, we can easily predict the learning disability of any child. 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. The Rough set theory approach and decision tree model classifier shows, its capability in discovering knowledge behind the LD identification procedure. The main contribution of this study is the comparison of Rough Set Theory and Decision Tree model for prediction of LD. In best of our knowledge, none of the study has conducted for prediction of LD. In this paper, we are considering an approach to handle learning disability database for the two data mining classification methods – Rough Sets Theory and Decision Trees for the prediction the learning disability in school age children. This study has been carried out on 513 real data sets with the attributes, which represents the symptoms of LD, takes binary values and more work need to be carried out on quantitative data, as that is an important part of any data set. Our future research work focus on, fuzzy-neuro methods, for finding the percentage of LD in each child.

8. REFERENCES

[1] Abraham, Ajith; Falcón, Rafael; Bello, Rafael (Eds.).2009. Rough Set Theory: A True Landmark in Data Analysis, Series: Studies in Computational Intelligence, Vol. 174, ISBN: 978-3-540-89920-4,

[2] Ashwin Kothari and Avinash Keskar, 2009. Paper on Rough Set Approach for Overall Performance Improvement of an Unsupervised ANN Based Pattern Classifier, Journal on Advanced Computational Intelligence and Intelligent Information, Vol. 13, No.4, 434-440

[3] Blackwell Synergy. 2006 Learning Disabilities Research Practices, Volume 22

[4] Chen R.S, Wu R.C., Chang C.C 2005. 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, SNPD/SAWN, Towson, MD, USA, 70-75

[5] Crealock Carol, Kronick Doreen. 1993. Children and Young People with Specific Learning Disabilities, Guides for Special Education, No. 9, UNESCO

[6] Fayyad U.M., 1996. From Data Mining to Knowledge Discovery:An Overview- Advances in Knowledge Discovery and Data Mining:-34, AAAI Press/MIT Press, ISBN 0-262-56097-6

[7] Frawley and Piaatetsky.1996. Shaping Knowledge Discovery in Database; An Overview, The AAAI/MIT press, Menlo Park

[8] Greco S., Matarazzo B, Slowinski R. 2000. Dealing with missing data in rough set analysis of multi-attribute and multi-criteria decision problem, Kluwer Academic Publishers, Boston Dordrecht, London

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

[10] Hameed Al-Qaheri, Aboul Ella Hassanien and Ajith Abraham. 2008. Discovering Stock Price Prediction Rules using Rough Sets

[11] Han Jiawei and Kamber Micheline, 2008. Data Mining-Concepts and Techniques, Second Edition, Morgan Kaufmann - Elsevier Publishers, ISBN : 978-1-55860-901-3

[12] Hsinchun Chen, Sherrilynne S. Fuller, Carol Friedman and William Hersh, 2005. Knowledge Discovery in Data Mining and Text Mining in Medical Informatics, Chapter 1, 3-34

[13] Iftikar U. Sikder, Toshinori Munakata, 2009. Application of rough set and decision tree for characterization of premonitory factors of low seismic activity, Expert system with applications, Elsevier, Vol 36, 102-110, available at www.sciencedirect.com

[14] Julie M. David, Pramod K.V, 2008. Paper on Prediction of Learning Disabilities in School Age Children using Data Mining Techniques. In: Proceedings of AICTE Sponsored National Conference on Recent Developments and Applications of Probability Theory, Random Process and Random Variables in Computer Science, T. Thrivikram, P. Nagabhushan, M.S. Samuel (eds), 139-146

[15] Julie M. David, Kannan Balakrishnan. 2009. Paper on Prediction of Frequent Signs of Learning Disabilities in School Age Children using Association Rules. In: Proceedings of the International Conference on Advanced Computing, ICAC 09, MacMillion Publishers India Ltd, NYC, ISBN 10:0230-63915-1, ISBN 13:978-0230-63915-7, 202-207

[16] Julie M. David, Kannan Balakrishnan, 2010. Paper on Prediction of Learning Disabilities in School Age Children using Decision Tree. In: Proceedings of the International Conference on Recent Trends in Network Communications-CCIS Vol. 90, Part 3, N. Meghanathan, Selma Boumerdassi, Nabendu Chaki, Dhinaharan Nagamalai (eds), Springer-Verlag Berlin Heidelberg, ISSN:1865-0929(print) 1865-0937 (online), ISBN 978-3-642-14492-9 (print) 978-3-642-14493-6 (online), DOI : 10.1007/978-3-642-14493-6_55, 533-542.

[17] Julie M. David, Kannan Balakrishnan, Oct.2010. Significance of Classification Techniques in Prediction of Learning Disabilities in School Age Children, International Journal of Artificial Intelligence & Applications (IJAIA), Vol 1, No. 4, DOI:10.5121/ijaia.2010.1409, 111-120

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

[19] Marcin Perzyk, Artur Soroczynski, 2010. Knowledge extraction tools for design and control of industrial process.

[20] Matteo Magnani. 2003. Technical report on Rough Set Theory for Knowledge Discovery in Data Bases

[21] Pawlak Z. 1982. Rough Sets, International Journal on Computers and Information Science, Vol. 11, 341-356

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

[23] Rod Paige, (Secretary). 2002. US Department of Education, Twenty-fourth Annual Report to Congress on the Implementation of the Individuals with disabilities Education Act-To Assure the Free Appropriate Public Education of all Children with Disabilities

[24] Sally Jo Cunningham and Geoffrey Holmes, 1999. Developing innovative applications in agricultural using data mining. In: Proceedings of the Southeast Asia Regional Computer Confederation Conference

[25] Stuart R., Peter N., 2009. Artificial Intelligence – A Modern approach, Pearson Prentice Hall

[26] Tan Pang-Ning, Steinbach Michael, Kumar Vipin, 2008. Introduction to Data Mining, Low Price Edition, Pearson Education, Inc., ISBN 978-81-317-1472-0

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

[28] Wang K., Aug. 2007. Applying data mining to manufacturing; The nature and implications, Journal of Intelligent Manufacturing, Vol. 18, No.4, 487-495

[29] Witten I.H, Frank Ibe. 2005. Data Mining Practical Machine Learning Tools and Techniques, Morgan Kaufmann Elsevier Publishers, 2nd Edition, ISBN : 13: 978-81-312-0050-6

BIOGRAPHIES

Julie M. David received her MCA degree from Bharathiyar University, Coimbatore, India in 2000, the M.Phil degree in Computer Science from Vinayaka Missions University, Salem, India in 2008 and is currently pursuing the Ph. D. degree in the research area of Data Mining from Cochin University of Science and Technology, Cochin, India. During 2000-2007, she was with Mahatma Gandhi University, Kottayam, India as Lecturer in the Department of Computer Applications. Currently she is working as Asst. Professor in the Department of Computer Applications with MES College, Aluva, Cochin, India. She has published papers in International Journals and International and National conference proceedings. Her research interests include Data Mining, Artificial Intelligence and Machine Learning. She is a member of International Association of Engineers and an International Reviewer of Elsevier Knowledge Based Systems.

Dr. Kannan Balakrishnan received his M.Sc and M. Phil degrees in Mathematics from University of Kerala, India, M. Tech degree in Computer and Information Science from Cochin University of Science & Technology (CUSAT), Cochin, India and Ph. D in Futures Studies from University of Kerala, India in 1982, 1983, 1988 and 2006 respectively. He is currently working with CUSAT, Cochin, India, as Associate Professor (Reader), in the Department of Computer Applications. He has visited Netherlands as part of a MHRD project on Computer Networks. Also he visited Slovenia as the co-investigator of Indo-Slovenian joint research project by Department of Science and Technology, Government of India. He has published several papers in International Journals and National and International conference proceedings. His present areas of interest are Graph Algorithms, Intelligent Systems, Image Processing, CBIR and Machine Translation. He is a reviewer of American Mathematical Reviews. He is a recognized Research Guide in the Faculties of Technology and Science in CUSAT, Cochin, India. He has served in many academic bodies of various universities in Kerala, India. Also currently he is a member of the Board of Studies of Cochin, Calicut and Kannur Universities in India. He is also a member of MIR labs India.