ISSN:0975-9646

Julie M. David et al, / (IJCSIT) International Journal of Computer Science and Information Technologies, Vol. 2 (2) , 2011, 829-835

Prediction of Learning Disabilities in School Age Children using SVM and Decision Tree

Julie M. David[1], Kannan Balakrishnan[2]

[1]*Dept. of Computer Applications, MES College, Marampilly, Aluva, Cochin- 683 107, India*

[2]*Dept. of Computer Applications, Cochin University of Science & Technology, Cochin - 682 022, India*

Abstract- **This paper highlights the prediction of Learning Disabilities (LD) in school-age children using two classification methods, Support Vector Machine (SVM) and Decision Tree (DT), with an emphasis on applications of data mining. About 10% of children enrolled in school have a learning disability. Learning disability prediction in school age children is a very complicated task because it tends to be identified in elementary school where there is no *one sign* to be identified. By using any of the two classification methods, SVM and DT, we can easily and accurately predict LD in any child. Also, we can determine the merits and demerits of these two classifiers and the best one can be selected for the use in the relevant field. In this study, Sequential Minimal Optimization (SMO) algorithm is used in performing SVM and J48 algorithm is used in constructing decision trees.**

Keywords- **Decision Tree, Hyper Plane, Learning Disability, Polykernel, Support Vector Machine**

I. INTRODUCTION

Databases are rich with hidden information, which can be used for intelligent decision making. 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 [22]. Knowledge Discovery in Databases (KDD) is the process of identifying useful information in data [18]. 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 [6]. 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 [5]. Diverse fields such as marketing, customer relationship management, engineering, medicine, crime analysis, expert prediction, web mining and mobile computing besides others utilize data mining [9]. 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 [1].

This paper presents the study of SVM and decision tree classifiers and shows how these ideas may be utilized for data mining. The SVM approach seems to be of fundamental importance to artificial intelligence and especially in the case of machine learning, knowledge acquisition, decision analysis, knowledge discovery from databases, expert systems, inductive reasoning and pattern recognition [24]. The SVM approach of data analysis has much important advantage and hence it has been successfully applied in many real life problems.

During the late 1970s and early 1980s, J. Ross Quinlan, a researcher in machine learning, developed decision tree algorithms known as ID3 [19]. 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.

Support vector machines and decision trees are probably the most frequently used tools for rule extraction from data [28, 3]. These classifiers are simple and easy to interpret by users whereas the SVM based methods seems to be their newer alternative. In this study, we have explored this relatively new family of learning methods called SVMs or, more generally Kernel Machines apart form DTs. To some extent, kernel machines give us the best of both worlds. That is, these methods use an efficient training algorithm and can represent complex, nonlinear functions [24]. In both cases, the practical aspects of application of these 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. SVM may require long computational times compared to DT. The classification algorithm is very important, particularly in construction of data mining system. Therefore, there are very little studies are available in the area of prediction of learning disabilities. The main purpose of the present paper is to show the important differences in performance of these two data mining methods in the prediction of learning disabilities affected in school age children.

This paper consists of four main parts. First, explanation about LD, second, description about SVM and DT, third comparison between the performance and results obtained for the classifiers SVM and decision tree type model using LD datasets and finally, in result analysis, the merits and demerits of these two classification methods, using LD data sets, is dealt with in detail.

II. LEARNING DISABILITY

Learning disability is a general term that describes specific kinds of learning problems. It is a neurological condition that affects a child's brain and impairs his ability to carry out one or many specific tasks [16]. Specific learning disabilities have been recognized in some countries for much of the 20th century, in other countries only in the latter half of the century, and yet not at all in other places. 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 [13]. 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. When a child 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 [12]. 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 [14]. 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 [21]. 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.

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 [1]. In fact, over half of all children who receive special education have a learning disability [4]. 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 [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.

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 advice it. Parental consent is needed before a child can be tested [16]. 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 [4]. 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 [13]. 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 [12]. Here for assessing the LD, we are using a checklist containing 16 symptoms of LD, which are the attributes in this study. The list of attributes is shown in Table I below.

TABLE I
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 Learning a Language
14	DLS	Difficulty Learning a Subject
15	STL	Slow To Learn
16	RG	Repeated a Grade

III. SUPPORT VECTOR MACHINES AND DECISION TREES

Vladimir Vapnik invented Support Vector Machine in 1979 [26,8]. Support Vector Machine algorithm is based on statistical learning theory. It is a new method for the classification of both linear and non-linear data. The basic idea behind the support vector machine is to map the original data into a feature space with high dimensionality through a non-linear mapping function and construct an optimal hyper plane in new space [20]. In the last few years, there has been a surge of interest in support vector machine [26,27]. SVM have empirically been shown to give good generalization performance on a wide variety of problems such as hand written character recognition, face detection and pedestrian detection [11]. SVM can be applied to both classification and regression. In the case of classification, an optimal hyper plane is found that separates the data into two classes, whereas in the case of regression a hyper plane is to be constructed that less close to as many points as possible. Separating the classes with a large margin minimizes a bound on the expected generalization error [11]. A minimum generalization error means that when new examples arrive for classification, the chance of making an error in the prediction based on the learned classifier should be minimum. Such a classifier is one, which achieve maximum separation margin between the classes. The two planes parallel to the classifier and which passes through one or more points in the data set are called bounding planes [23]. Support Vector Machines select a small number of critical boundary instances called support vectors from each class and build a linear discriminant function that separates them as widely as possible [29]. The points in the dataset falling on the bounding planes are called support vectors. SVM algorithm transforms the original data in a higher dimension, from where it can find a hyper plane for separation of the data using essential training tuples called support vectors [7]. These points play a crucial role in the theory and hence the name Support Vector Machines. Machine means algorithm.

If the training vectors are separated without errors by an optimal hyper plane, the expected error rate on a test sample is bounded by the ratio of the expectation of the support vectors to the number of training vectors. Since this ratio is independent of the dimension of the problem, if one can find a small set of support vectors, good generalisation is guaranteed. In the case, one may simply minimise the number of misclassification whilst maximising the margin with respect to the correctly classified instances. In such a case, it is said that the SVM training algorithm allows a training error [23]. There may be another situation; the points are clustered such that the two classes are not linearly separable. It may have to tolerate large training error. In such cases, we prefer nonlinear mapping of data into some higher dimensional space called feature space, where it is linearly separable. In order to distinguish between these two spaces, the original space of data point is called input space [23]. The hyper plane in feature space corresponds to a highly nonlinear separating surface in the original input space. Hence the classifier is called nonlinear classifier [2].

However, the use of SVM is still limited to a small group of researchers. One possible reason is that training algorithm for SVM is slow especially for large problems. Another explanation is that, SVM training algorithm is complex, subtle and difficult for an average engineer to implement [11].

Whereas 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 [7]. The topmost node in a tree is the root node. Decision trees are powerful and popular tool for classification and prediction [14]. 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. The learning and classification step of decision tree are simple and fast [14]. 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 [25].

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 [24]. 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 [24].

IV. PERFORMANCE COMPARISON OF CLASSIFIERS AND RESULTS OBTAINED

In this paper, we use the learning algorithms, Sequential Minimal Optimization (SMO) and J48, for the prediction of LDs in the classifiers SVM and DT respectively. The performances of these classifiers along with the results obtained are compared. SMO is conceptually simple, easy to implement and generally faster. SVMs belong to the class of supervised learning algorithms in which the learning machine is given a set of examples with the associated labels as in the case of decision trees, the examples are in the form of attribute vectors [23].

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 sometimes it can give wrong predictions when inconsistent data are present. In the case of prediction of LD, wrong prediction result will make a large problem. So we consider the solution for recovering that problem and use the simplicity of decision tree structure. In this study, we are used 513 real data sets collected from various schools in and around Cochin, India, for learning disability prediction in children using both these classifiers.

A. Results of SVM

Sequential minimal optimization algorithm is used for training a support vector classifier. This implementation globally replaces all missing values and transforms nominal attributes into binary ones. It also normalizes all attributes by defaults. In this study, the SMO algorithm which uses the polykernel, correctly classified 97.86% instances from the real data sets with a complexity parameter of 1. The accuracy of SVM is given in Table II below.

TABLE II
ACCURACY OF SVM

TP Rate	**FP Rate**	**Pre-cision**	**Recall**	**F Mea-sure**	**ROC Area**	**Class**
0.990	0.030	0.981	0.990	0.986	0.980	T
0.979	0.016	0.974	0.979	0.977	0.978	F
Correctly Classified Instances					502 Nos.	97.86%
Incorrectly Classified Instances					11 Nos.	2.14 %
Time taken to build a model						2.78Sec

B. Results of Decision Tree

We used J48 algorithm in weka, a machine-learning workbench, which include a framework in the form of Java class library [10]. 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, the J48 algorithm in weka, used for constructing the tree, correctly classified 97.47% instances from the real data sets. The accuracy of decision tree is given in Table III

TABLE III
ACCURACY OF DECISION TREES

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

For the construction of decision tree, the selection of attribute is very important. The inconsistent data may lead to false attribute selection in the case of decision tree. In the case of SVM, alternative methods are available in handling missing data. In this paper, we are using information gain as the attribute selection method in decision tree. But the inconsistency of the data leads to the false determination of attribute. The input values considered as the symptoms of LD. So the SVM and decision trees 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 decision tree generated is shown in Fig. 1 below.

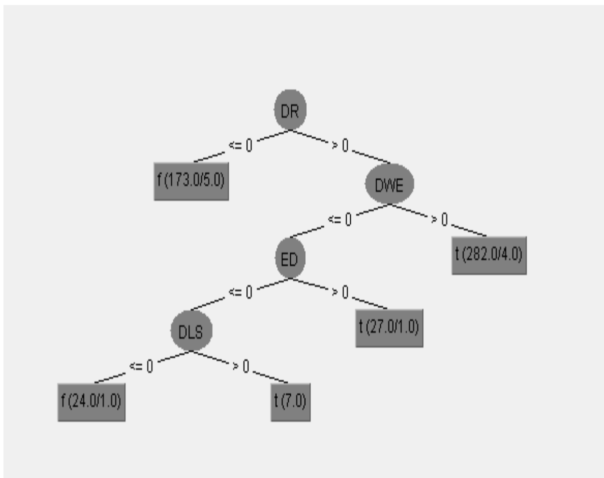

Fig. 1 Decision Tree

V. RESULT ANALYSIS

We can see that, both methods, viz. SVM and DT, 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. In SVM, the classification is mainly based on the type of kernel choose. Here we are using the polykernel for attribute evaluation. The confusion matrix in respect of SVM and DT is compared as shown in Table IV below.

TABLE IV
COMPARISON OF CONFUSION MATRICES

Support Vector Machine			**Decision Tree**		
a	b	←classified as	a	b	←classified as
312	3	a=T	310	5	a=T
4	190	b=F	4	190	b=F

The wrong predictions obtained from decision trees for 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. The accuracy level of the decision trees shows little lower accuracy compared to SVM. The cost curves of SVM and DT models are shown in Fig. 2 below.

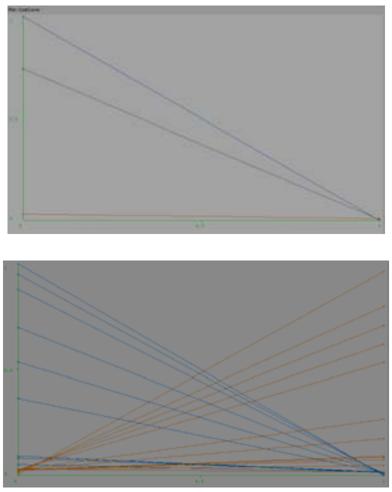

Fig. 2 Cost curve of SVM and DT models

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. We have used 513 real data sets, for the study.

In this study, we used SMO algorithm in SVM and J48 algorithm for constructing DT, for prediction of LD in children. From the comparison of results, we have noticed that SVM with SMO algorithm has a number of advantages over DT with J48 algorithm for solving such nature of problems. In the case of 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 SVM, the rules formulated will never influenced by any such incomplete datasets or attributes. Hence, it is found that LD can accurately be predicted by using both methods. If SVM is comparing with decision trees, the data or the output of SVM is very complex. However, the output of decision tree is categorical.

This study reveals that, out of the 513 real data sets, the SVM correctly classifies 502 instances in 2.78 seconds whereas DT correctly classifies 500 instances in 0.08 seconds. This shows that though the difference is very little, SVM is more suitable in getting accurate results in prediction of LD. The accuracy of the classifiers can be determined by ROC curve. The area under ROC curve in both cases is nearer to 1, which means the accuracy of both classifiers is found to good.

VI. CONCLUSIONS AND FUTURE RESEARCH

This paper highlights the prediction of Learning Disabilities in school-age children using two classification methods, Support Vector Machine (SVM) and Decision Tree (DT), with an emphasis on applications of data mining. By using any of these two classification methods, we can easily and accurately predict LD in any child. Both the methods are very effective for the prediction. However, the SVM gives more accurate result compared to DT, even though the time taken to build the model is much higher. The wrong predictions obtained from decision trees for inconsistent data sets can be lead to a limited accuracy of decision tree models. The result of this study indicates that, the rules system represented by the decision trees may be significantly incorrect for inconsistent data. 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 SVM may require long computational times. Hence by this study, we can determine the merits and demerits of these two classifiers while using in the relevant field.

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 using any of these classifiers, we can easily predict the learning disability of any child. SVM 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 performances of SVM and decision tree models with an emphasis on the results obtained while applying in 513 real data sets for prediction of LD in children. 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 studying the merits and demerits of the two data mining classification methods – Support Vector Machine and Decision Trees for the prediction of learning disability in school age children. This study has been carried out on more than 500 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. In comparison with our another study using rough sets and decision trees [17], in this field, we found that rough sets is more suitable in attribute selection while decision tree is suitable in classification. Our future work focuses on fuzzy sets and near sets for predicting LD.

REFERENCES

[1] Blackwell Synergy, *Learning Disabilities & Research Practice*, Volume 22, 2007

[2] Boser B.E., Guyon I.M., Vapnik V., “A Training Algorithm for optimal Margin Classifiers”, Fifth Annual Workshop on Computational Learning Theory, ACM, 1992

[3] 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,* SNPD/SAWN 2005, Towson, MD, USA, 2005, 70-75

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

[5] Fayyad U.M., *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, 1996

[6] Frawley and Piaatetsky, *Shaping Knowledge Discovery in Database; an Overview*, The AAAI/MIT press, Menlo Park, 1996

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

[8] Haykin S, *Neural Networks-A Comprehensive Foundation*, Prentice Hall, 1999

[9] Hsinchun Chen, Sherrilynne S. Fuller, Carol Friedman and William Hersh. “Knowledge Discovery in Data Mining and Text Mining in Medical Informatics”, Chapter 1, 2005, pp3-34

[10] Iftikar U. Sikder, Toshinori Munakata, “Application of rough set and decision tree for characterization of premonitory factors of low seismic activity, Expert system with applications”, Elsevier, 36, 2009, pp 102-110, Available:www.sciencedirect.com

[11] John C. Platt, Sequential Minimal Optimisation – “A Fast Algorithm for Training Support Vector Machines”, Technical Report MSR-TR-98-14, April 21, 1998

[12] Julie M. David, Pramod K.V, “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), 2008, pp 139-146

[13] Julie M. David, Kannan Balakrishnan, “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 2009,* MacMillion Publishers India Ltd., NYC, ISBN 10:0230-63915-1, ISBN 13:978-0230-63915-7, 2009, pp. 202–207

[14] Julie M. David, Kannan Balakrishnan, “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-* Communication in Computer and Information Science, 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, 2010, pp 533-542

[15] Julie M. David, Kannan Balakrishnan, “Prediction of Key Symptoms of Learning Disabilities in School Age Children using Rough Sets” International Journal of Computer and Electrical Engineers (IJCEE), Vol. 3, No. 1, Feb.2011, pp 163-169.

[16] Julie M. David, Kannan Balakrishnan, “Significance of Classification Techniques in Prediction of Learning Disabilities” International Journal of Artificial Intelligence & Applications (IJAIA), Vol 1, No. 4, DOI:10.5121/ijaia.2010.1409, Oct.2010, pp111-120.

[17] Julie M. David, Kannan Balakrishnan, “Machine Learning Approach for Prediction of Learning Disabilities in School Age Children”, International Journal of Computer Applications(IJCA), ISSN-0975-8887, Nov.2010, pp 7-14

[18] Matteo Magnani. “Technical Report on Rough Set Theory for Knowledge Discovery in Data Bases”, 2003

[19] Quinlan J.R., Induction on Decision Trees, Machine Learning, 1(1): 81-106, 1986

[20] Radhika Y, Shashi M., “Atmospheric Temperature Prediction using Support Vector Machines”, *International Journal of Computer Theory and Engineering*, Vol. 1, No.1, April 2009, 1793-8201 55-58

[21] Rod Paige (Secretary) 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”, 2002

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

[23] Soman K.P., Loganathan R., Ajay V, *Machine Learning with SVM and other Kernel Methods*, New Delhi, PHI Learning Pvt. Ltd, ISBN-978-81-203-3435-9, 2009

[24] Stuart Russell, Peter Norvig, *Artificial Intelligence – A Modern Approach*, Pearson Prentice Hall, 2009

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

[26] Vapnik V, *Estimation of Dependences Based on Empirical Data*, Springer-Verlag, 1982

[27] Vapnik V, *The Nature of Statistical Learning theory*, Springer-Verlag, 1995

[28] Wang K., *Applying data mining to manufacturing; The nature and implications*, J. Intell Manuf, vol 18, no.4, Aug 2007, 487-495

[29] Witten Ian H, Frank ibe, Data Mining – *Practical Machine Learning Tools and Techniques*, Morgan Kaufmann - Elsevier Publishers, ISBN: 13: 978-81-312-0050-6, 2005

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 several 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 a 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, 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 Cochin University of Science & Technology, 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 the Cochin University of Science and Technology, 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.