



International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 2, Issue 7, July 2014

Comparative Study of Decision Trees and Rough Sets for the Prediction of Learning Disabilities in School-Age Children

Dr. Julie M. David¹, Dr. Kannan Balakrishnan²

¹Dept. of Computer Applications, MES College, Marampally, Aluva Cochin- 683 107, India

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

ABSTRACT: This paper highlights the study of two classification methods, Rough Sets Theory (RST) 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 classification methods we can easily and accurately predict LD in any child. Also, we can determine the best classification method. In this study, rule mining is performed using the algorithms LEM1 in rough sets and J48 in construction of decision trees. From this study, it is concluded that, the performance of decision trees may be considerably poorer in several important aspects compared to that of rough sets theory. It is found that, for selection of attributes, RST is very useful especially in the case of inconsistent data.

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

I. INTRODUCTION

This paper presents the comparative study for rough sets and decision trees and shows how these ideas may be utilized for data mining. During the late 1970s and early 1980s, J. Ross Quinlan, a researcher in machine learning developed a decision tree algorithms known as ID3 [8]. 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.

For prediction of LD, decision trees are probably the most frequently used tools for rule extraction from data whereas the rough sets based methods seems to be their newer alternative. In both cases, the algorithms are simple and easy to interpret by users. There are very little comparative studies are available. The purpose of the present paper is to show the important differences in performance of two data mining methods for the prediction of LD in children. The rough set 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 [2].

II. RELATED WORK

Learning disability is a general term that describes specific kinds of learning problems. Learning disabilities are formally defined in many ways in many countries. The most frequent clause used in determining whether a child has a learning disability is the difference between areas of functioning. 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 [5]. 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 [5]. 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 [9]. With the right help, children with LD can and do learn successfully. 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.

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 2, Issue 7, July 2014

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

III. PROPOSED ALGORITHM

A. Design Considerations:

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 [3]. The topmost node in a tree is the root node. 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. 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. 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.

Rough set theory is a new intelligent mathematical tool introduced by Z. Pawlak in 1982[7]. 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 [6]. A rough set is an approximation tool that works well when in environments heavy with inconsistency and ambiguity in data or involving missing data [1].

B. Description of the Proposed Algorithm:

We used J48 algorithm in weka, a machine learning workbench, which include a framework in the form of Java class library [4]. 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 75% instances from the data sets using weka. The obtained rules are summarised below.

- | | |
|--|--|
| R1: (DR=N, DA=N) => (LD, N) (1) | R4: (DR=Y, DBA=N, DLS=N, DSS=N) => (LD, N) (4) |
| R2: (DR=N, DA=Y, DH=Y) => (LD, Y) (2) | R5: (DR=Y, DBA=N, DLS=Y) => (LD, Y) (5) |
| R3: (DR=N, DA=Y, DHA=N) => (LD, N) (3) | R6: (DR=Y, DBA=N, DLS=N, DS=Y) => (LD, Y) (6) |

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 2, Issue 7, July 2014

$$R7: (DR=Y, DBA=Y) \Rightarrow (LD, Y) \quad (7)$$

The rough set application development consist four steps. The first step is the development of decision table. Decision table include 100 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 a relatively straight forward 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 the following six rules that predict the learning disability.

$$\begin{array}{ll}
 R1:(DR, Y)(DS, Y)(DH, N)DWE, Y) = (LD, Y) & (1) \\
 R2:(DH, N) (DWE, N) = (LD, N) & (2) \\
 R3:(DH, Y) (DWE, Y) = (LD, Y) & (3) \\
 R4:(DH, N) (DWE, Y) = (LD, Y) & (4) \\
 R5:(DWE, Y) = (LD, Y) & (5) \\
 R6:(DWE, N) = (LD, N) & (6)
 \end{array}$$

IV. SIMULATION RESULTS

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, is based on elimination of redundant attribute in a decision table. The focus is to identify minimal set of attribute that preserve the indiscernibility relation.

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 2 below. If the same rules applied on the inconsistent data, the confidence of the rules is reducing to a poor level, which is also shown in Table 2. The confidence of these rules, based on RST, shows a higher performance, as shown in the same table, as compared to DT with consistent data.

Table 2. Confidence of rules

Sl. No.	Rules	Confidence		
		DT with consistent data	DT with inconsistent data	RST
1	R1	73%	53%	80%
2	R2	72%	52%	80%
3	R3	71%	41%	60%
4	R4	70%	30%	85%
5	R5	71%	61%	80%
6	R6	73%	65%	90%
7	R7	70%	50%	-

In this study, we can see that RST is more suitable and accurate for selecting attributes. For the construction of decision tree the selection of attribute is very important. The rough set theory has been used for selecting attributes,



International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 2, Issue 7, July 2014

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.

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 comparative study indicates that the rules system represented by the decision trees may be significantly in correct 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.

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.

In this study, we are used LEM1 algorithm in RST, for rule mining and J48 algorithm for constructing DT, for prediction of LD in children. From the comparison of results, we have noticed that RST with LEM1 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 it is comparing with decision trees, the data or the output of decision tree is very complex. Another thing is that the output of decision tree is categorical.

V. CONCLUSION AND FUTURE WORK

This paper highlights the comparison between the Decision Tree and Rough Set Theory to predict the learning disabilities in school age children. In Rough Set Theory, LEM1 algorithm is used for the rule generation and J48 algorithm is used for the construction of decision tree. The extracted rules in both the methods are very effective for the prediction. 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 comparative 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 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. This study has been carried out on more than 100 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.

REFERENCES.

1. Ashwin Kothari and Avinash Keskar. Paper on Rough Set Approach for Overall Performance Improvement of an Unsupervised ANN-Based Pattern Classifier, 2009
2. Grzymala-Busse JW. Knowledge Acquisition under Uncertainty-A Rough Set Approach. Journal of Intelligent & Robotic Systems, 1988, 1: 3-16
3. Han Jiawei and Kamber Micheline, : Data Mining-Concepts and Techniques, Second Edition, Morgan Kaufmann - Elsevier Publishers, ISBN : 978-1-55860-901-3, 2008
4. 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, 102-110



ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 2, Issue 7, July 2014

5. Julie M. David, Kannan Balakrishnan. "Paper on Prediction of Frequent Signs of Learning Disabilities in School Age Children using Association Rules", Proceedings of the International Conference on Advanced Computing, ICAC 09, MacMillan Publishers India Ltd, ISBN 10:0230-63915-1, ISBN 13:978-0230-63915-7, 2009, 202-207
6. Matteo Magnani. Technical report on Rough Set Theory for Knowledge Discovery in Data Bases, 2003
7. Pawlak Z. Rough Sets. Int. J. Computers and Information Sci., Vol 11, 1982, 341-356
8. Quinlan J.R., Induction on decision trees, Machine learning, 1(1):81-106,1986
9. 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