The Result Oriented Process for Students Based On Distributed Data Mining

Pallamreddy.venkatasubbareddy[1]

Associate Professor, CSE Dept
QIS College of Engg & Technology
Ongole, India

Vuda Sreenivasarao[2]

Professor & HOD, CSE Dept
St.Mary's College of Engg & Technology,
Hyderabad, India.

Abstract - **The student result oriented learning process evaluation system is an essential tool and approach for monitoring and controlling the quality of learning process. From the perspective of data analysis, this paper conducts a research on student result oriented learning process evaluation system based on distributed data mining and decision tree algorithm. Data mining technology has emerged as a means for identifying patterns and trends from large quantities of data. It is aimed at putting forward a rule-discovery approach suitable for the student learning result evaluation and applying it into practice so as to improve learning evaluation of communication skills and finally better serve learning practicing.**

Key words: Learning result evaluation system; distributed data mining; decision tree.

I. INTRODUCTION

With the accelerating development of society and the well-known knowledge explosion in modern times, learning is taking on a more important role in the development of our civilization. Learning is an individual behavior as well as a social phenomenon. For a long time, people limited learning to the transfer of culture and knowledge; research on learning was confined to the fields of education research within the existing traditions of classroom learning. To understand learning within the new context of the 21st century, we need professionals from psychology, sociology, bran science, computer science, economics, to name a few. We must extend out understanding about human learning from macro levels to micro levels and from history to current conditions. At present, the most urgent need is to synthesize all the findings on human learning and integrate them into a united framework to guide the practice of learners. A research on result oriented learning process evaluation from the perspectives of philosophy and culture has emerged as a major challenge to educators, and invent a new and integrated learning culture.

It is a difficult task to deeply investigate and successfully develop models for evaluating learning efforts with the combination of theory and practice. University goals and outcomes clearly relate to "promoting learning through effective undergraduate and graduate teaching, scholarship, and research in service to University." Student learning is addressed in some following university goals and outcomes related to the development of overall student knowledge, skill, and dispositions. Collections of randomly selected student work are examined and assessed by small groups of faculty teaching courses within some general education categories. Education reform for the 21st century has generated various models of learning result evaluation that have emerged over time.

The assessment of student learning is an essential component of University's efforts in evaluating overall institutional effectiveness. The value of research on student learning result evaluation system is to help teachers and students surpass the ivory-towered and alienating traditional classroom teaching model, and face the rapidly developing real-life environment and the ill-structured learning environment and adapt to current teaching realities. Distributed Data Mining (DDM) aims at extraction useful pattern from distributed heterogeneous data bases in order, for example, to compose them within a distributed knowledge base and use for the purposes of decision making. A lot of modern applications fall into the category of systems that need DDM supporting distributed decision making. Applications can be of different natures and from different scopes, for example, data and information fusion for situational awareness; scientific data mining in order to compose the results of diverse experiments and design a model of a phenomena, intrusion detection, analysis, prognosis and handling of natural and man-caused disaster to prevent their catastrophic development, Web mining ,etc. From practical point of view, DDM is of great concern and ultimate urgency.

The remaining sections of the paper are organized as follows. In Section II we describe the distributed data mining. . In Section III we describe decision tree algorithm. In Section IV we describe Student Learning Result Evaluation System and verify the efficiency of this system. Section V concludes the paper.

II. DISTRIBUTED DATA MINING

Data mining technology has emerged as a means for identifying patterns and trends from large quantities of data. Data mining and data warehousing go hand-in-hand: most tools operate on a principal of gathering all data into a central site, then running an algorithm against that data (Figure 1). There are a number of applications that are infeasible under such a methodology, leading to a need for distributed data mining.
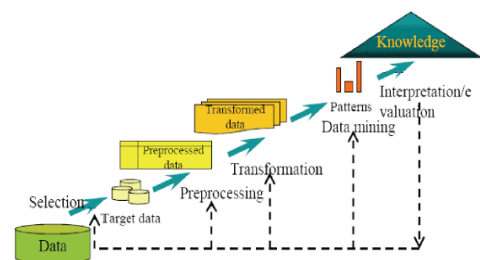

Fig 1 the transaction from raw data to valuable knowledge.

Distributed Data Mining Distributed data mining (DDM) considers data mining in this broader context. As shown in figure(2), objective of DDM is to perform the data mining operations based on the type and availability of the distributed resources. It may choose to download the data sets to a single site and perform the data mining operations at a central location.
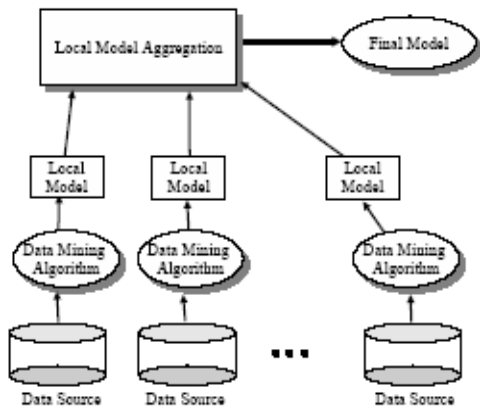

Fig (2). Distributed Data mining Frame work

Data mining is a powerful new technology with great potential to help companies focus on the most important information in the data they have collected about the behavior of their customers and potential customers. Data mining involves the use of sophisticated data analysis tools to discover previously unknown, valid patterns and relationships in large data set. These tools can include statistical models, mathematical algorithm and machine learning methods. It discovers information within the data that queries and reports can't effectively reveal.

III. DECISION TREE ALGORITHM

A decision tree is a classification scheme which generates a tree and a set of rules, representing the models of deferent classes, from a given data set. The set of records available for developing classification methods is generally divided in to two disjoint subsets-a training set and a test set. The decision tree approach is most useful in classification problems, with this technique, a tree is constructed to modal the classification process.

Decision tree is a decision support tool in the field of data mining that uses a tree-like graph or model of decisions and their possible consequences, including chance event outcomes, resource costs, and utility . Decision trees are commonly used in operations research, specifically in decision analysis, to help identify a strategy most likely to reach a goal. Another use of decision trees is as a descriptive means for calculating conditional probabilities. Each node in a decision tree represents a feature in an instance to be classified, and each branch represents a value that the node can assume. Instances are classified starting at the root node and sorted based on their feature values.

A .Definition:

Given a data base $D = \{t_{1,\ldots\ldots\ldots} tn\}$ where

$ti = \{ti1,\ldots\ldots tih\}$ and the data base schema contains the following attributes $\{A1,A2,A3,\ldots\ldots Ah\}$. Also given is a set of classes $C=\{C1,C2,C3\ldots\ldots Cm\}$.

A decision tree (DT) or classification tree is a tree associated with D that has the following properties:-

Each internal node is labeled with an attribute, Ai

Each arc is labeled with a predicate that can be applied to the attribute associated with the parent.

Each leaf node is labeled with a class, Cj

The basic algorithm for decision tree induction is a greedy algorithm that constructs decision trees in a top-down recursive divide-and-conquer manner. The algorithm is shown in fig.3.

create a node N

if

    samples are all of the same class C

then

    return N as a leaf node labeled with the class C

if

    attribute-list is empty

then

    return N as a leaf node labeled with the most common class

    in samples

    label node N with test-attribute

for each known value a_i of test-attribute

    grow a branch from node N for the condition test-attribute=a_i

    let s_i be the set of samples for which test-attribute= a_i

if

    s_i is empty

then

    attach a leaf labeled with the most common class in

    samples

else

    Attach the node returned by Generate_ decision_ tree

(s$_i$, attribute-list_ test-attribute)

Fig. 3   Algorithm for decision tree .

IV. RESULT ORIENTED LEARNING EVALUATION SYSTEM FOR STUDENT

Result Oriented Learning Evaluation System for Student (ROLES) is composed of five parts. Data Collection Module collects data. Data Processing Module converts rude data to the regular mode. Data Analysis Module use Decision Tree to compute regular data and give the corresponding result. Graph Visualizing Module is used to show data analysis result through graphic mode. Database is used to stored rude

data, regular data and data analysis result. The structure of is shown in fig.4.

Fig.5 shows the Decision Tree used in ROLES. Each internal node tests an attribute, each branch corresponds to attribute value, each leaf node assigns a classification.

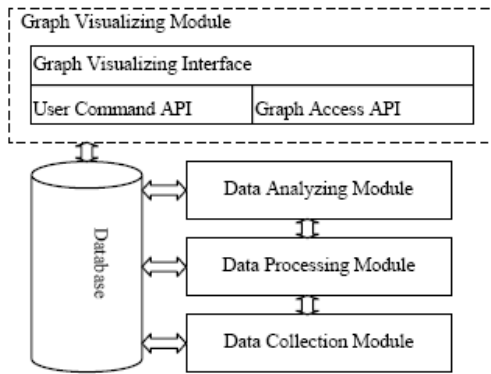Fig.4 shows the accuracy of decision tree learning.
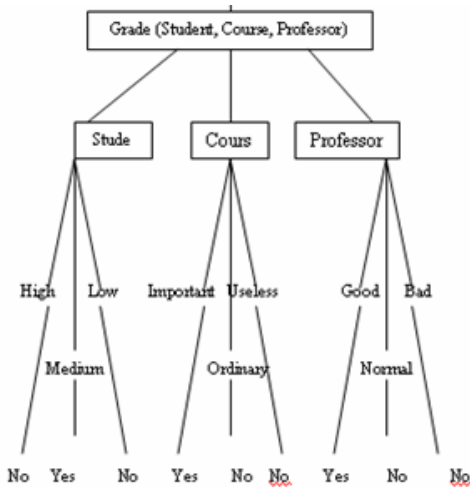

Fig.4: Structure of ROLES.

Figure No- 5 Decision tree used in ROLES.

No	Student	Course	Professor	Grade	Result
1	Ramu	DW	Srinivas	70	No
2	Ramya	DB	Swamy	69	No
3	Srinu	MS	Trirumala	95	Yes
4	Smith	DBMS	Raghavendra	97	Yes
....	....	....	....	....	....
n	Jone	CO	Prathibha	65	No

Table 1: Form of training examples

Table 1 shows the form of training data. 1000 student score records are used for training.
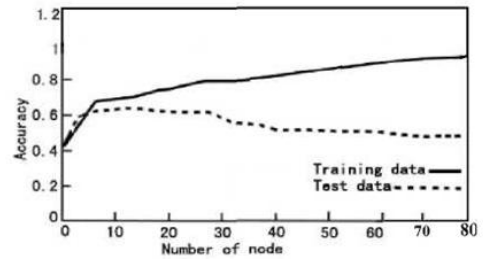

Figure 6: Accuracy of Decision tree learning.

V. CONCLUSION

The Student result oriented learning process system for assess learning environments is an image of what makes sense today. As time goes on, new features will be added and others dropped. With this model in practice, student learning can become more energetic, more interesting, more challenging, and more suited to the times.

REFERENCES

[1] Sally A. Gauci, Arianne M. Dantas, David A. Williams, and Robert E. Kemms, "Promoting Student-Centered Active Learning in Lectures with a Personal Response System," Advances in Physiology Education, vol. 33 (March 2009), pp. 60–71.

[2] Margie Martyn, "Clickers in the Classroom: An Active Learning Approach," EDUCAUSE Quarterly, vol. 30, no. 2, (April–June 2007), pp. 71–74.

[3] Y. Lindell and B. Pinkas Secure Multiparty Computation for Privacy-Preserving Data Mining Journal of Privacy and Confidentiality, Vol. 1, No. 1, pp. 59-98, 2009.

[4] T.C. Fu and C.L Lui, "Agent-oriented Network Intrusion Detection System Using Data Mining Approaches," International Journal on Agent-Oriented Software Engineering, Vol. 1, No. 2, pp. 158-174, 2007.

[5] Sang,X.M.: On the Edge of Learning Theory and Practice in the Digital Age. China Central Broadcasting and TV University Press, Beijing, 2000.

[6] R.A.Burnstern and L.M. Lederman, "Comparison of Different Commercial Wireless Keypad Systems", The Physics Teacher, Vol.41, Issure 5, 2003, pp.272-275.

[7] D.English, "Audiences Talk Back: Response Systems Fill Your Meeting Media with Instant Data", AV Video Multimedia Producer, Vol.25, No.12, 2003,pp.22-24.

AUTHOR PROFILE

PALLAMREDDY.VENKATA SUBBA REDDY received his M.Tech degree in Computer Science & Engg from the Satyabama University, in 2007 and Master of Science in Information Technology from Madras University with First class in may 2003 .Currently working as Assoc. Professor in CSE Dept in QIS College of Engg & Technology, Ongole. His main research interests are Data Mining and Network Security. He has got 7 years of teaching experience. He is currently pursuing the PhD degree in CSE Depart in Bharath University, Chennai.

VUDA SREENIVASARAO received his M.Tech degree in Computer Science & Engg from the Satyabama University, in 2007.Currently working as Professor & Head in the Department of CSE & IT at St.Mary's college of Engineering & Technology, Hyderabad, India.. His main research interests are Data Mining, Network Security, and Artificial Intelligence. He has got 10years of teaching experience .He has published 20 research papers in various international journals. He is a life member of various professional societies like MIACSIT, MISTE.