International Journal of Soft Computing and Engineering (IJSCE)
ISSN: 2231-2307, Volume-2, Issue-6, January 2013

Rule Mining Framework for Students Performance Evaluation

P. Ajith, B. Tejaswi, M.S.S.Sai

Abstract: Academic Data Mining used many techniques such as Decision Trees, Neural Networks, Naïve Bayes, K- Nearest neighbor, and many others. Using these techniques many kinds of knowledge can be discovered such as association rules, classifications and clustering. The discovered knowledge can be used for prediction and analysis purposes of student patterns. Prior approaches used decision tree classifications optimized with ID3 algorithms to obtain such patterns. Among sets of items in transaction databases, Association Rules aims at discovering implicative tendencies that can be valuable information for the decision-maker which is absent in tree based classifications. So we propose a new interactive approach to prune and filter discovered rules. First, we propose to integrate user knowledge in the post processing task. Second, we propose a Rule Schema formalism extending the specifications to obtain association rules from knowledge base. Furthermore, an interactive framework is designed to assist the user throughout the analyzing task. Applying our new approach to discover the likelihood of students deviations / requiring special attention is organized and efficient providing more insight by considering more information. Compared to tree based classifications the results are better to understand and can be applied to real time use. An implementation of the proposed system validates our claim.

Index Terms—Association Rules, Knowledge base, Prediction, and Rule Schema.

I. INTRODUCTION

The topic of explanation and prediction of academic performance is widely researched. Data Mining Techniques is the promising methodology to extract valuable information in this objective. The data collected from different applications require proper method of extracting knowledge from large repositories for better decision making. Knowledge discovery in databases (KDD), often called data mining, aims at the discovery of useful information from large collections of data. In this perspective, Data Mining can analyze relevant information results and produce different perspectives to understand more about the students' activities so as to customize the course for student learning.

Data mining task is used in computer science and information technology aspects such as online learning and collaborative learning to facilitate students learning. Results are satisfactory because the existing technology aids and addresses the aspects of automated learning, practicing and evaluations of an academic cycle. They facilitate to understand/monitor students performance based on that 'moment' scores. At that time, there is no perfect usage of data mining techniques to facilitate Students Learning. So a better system is required to monitor and analyze student's performance based on a knowledge base constructed from automated learning, practicing and evaluations of the academic cycle.

For optimally analyzing the student performance, the classification task is used on student database to predict the students division on the basis of previous database. As there are many approaches that are used for data classification, the decision tree method is used here. Information's like Attendance, Class test, Seminar and Assignment marks were collected from the student's previous database, to predict the performance at the end of the semester. This study will help the students and the teachers to improve the division of the student. This study will also work to identify those students which needed special attention to reduce fail percentages and taking appropriate action for the upcoming evaluations. This kind live performance monitoring and counter measures before the big evaluation definitely helps to improve students' performance.

In this paper, we use Association Rules Instead of tree based classification. A result of a tree based classification is complicated to understand and depends on the technical competency of the decision maker. Among sets of items in transaction databases, Association Rules aims at discovering implicative tendencies that can be valuable information for the decision-maker which is absent in tree based classifications. So we propose a new interactive approach to prune and filter discovered rules. First, we propose to integrate user knowledge in the post processing task. Second, we propose a Rule Schema formalism extending the specifications to obtain association rules from knowledge base. The decision maker need not have any technical competency to understand the results. Compared to tree based classifications the results are better to understand and can be applied to real time use.

II. DATA MINING TECHNIQUES

A. The K-means algorithm

The k-means algorithm is a simple iterative method to partition a given dataset into a userspecified number of clusters, k. The algorithm is initialized by picking k points as the initial k cluster representatives or "centroids". Techniques for selecting these initial seeds includes sampling at random from the dataset, setting them as the solution of clustering a small subset of the data or perturbing the global mean of the data k times. Then the algorithm iterates between two steps till convergence:
Step 1: Data Assignment. Each data point is assigned to its *closest* centroid, with ties broken arbitrarily. This results in a partitioning of the data.
Step 2: Relocation of "means". Each cluster representative is relocated to the center (mean) of all data points assigned to it. If the data points come with a probability measure (weights), then the relocation is to the expectations (weighted mean) of the data partitions.

Manuscript received on January, 2013.

P. Ajith, Research Scholar, KL University, Associate Prof. SRIJI College, Ongole, Andhra Pradesh, India.

B. Tejaswi, Asst. Prof. Department of CSE,KKR & KSR Institute of Technology & Sciences, Guntur, Andhra Pradesh, India.

M.S.S.Sai, Professor. Department of CSE, KKR & KSR Institute of Technology & Sciences, Guntur, Andhra Pradesh, India.

B. The Apriori algorithm

Apriori is a seminal algorithm for finding frequent itemsets using candidate generation. It is characterized as a level-wise complete search algorithm using anti-monotonicity of itemsets, "if an itemset is not frequent, any of its superset is never frequent". Apriori first scans the database and searches for frequent itemsets of size 1 by accumulating the count for each item and collecting those that satisfy the minimum support requirement. It then iterates on the following three steps and extracts all the frequent itemsets.

1. Generate $Ck+1$, candidates of frequent itemsets of size $k+1$, from the frequent itemsets of size k.
2. Scan the database and calculate the support of each candidate of frequent itemsets.
3. Add those itemsets that satisfies the minimum support requirement to $Fk+1$.

C. kNN: k- nearest neighbor classification

A more sophisticated approach, *k*-nearest neighbor (*k*NN) classification, finds a group of *k* objects in the training set that are closest to the test object, and bases the assignment of a label on the predominance of a particular class in this neighborhood. There are three key elements of this approach: a set of labeled objects, e.g., a set of stored records, a distance or similarity metric to compute distance between objects, and the value of *k*, the number of nearest neighbors. To classify an unlabeled object, the distance of this object to the labeled objects is computed, its *k*-nearest neighbors are identified, and the class labels of these nearest neighbors are then used to determine the class label of the object.

D. Naïve Bayes

Naïve Bayes method is important for several reasons. It is very easy to construct, not needing any complicated iterative parameter estimation schemes. This means it may be readily applied to huge data sets. It is easy to interpret, so users unskilled in classifier technology can understand why it is making the classification it makes. And finally, it often does surprisingly well: it may not be the best possible classifier in any particular application, but it can usually be relied on to be robust and to do quite well. Here, the initial set of objects are used with known class memberships (the training set) to construct a score such that larger scores are associated with class 1 objects (say) and smaller scores with class 0 objects. Classification is then achieved by comparing this score with a threshold, *t*.

E. CART

The CART decision tree is a binary recursive partitioning procedure capable of processing continuous and nominal attributes both as targets and predictors. Data are handled in their raw form; no binning is required or recommended. Trees are grown to a maximal size without the use of a stopping rule and then pruned back to the root via cost-complexity pruning. The next split to be pruned is the one contributing least to the overall performance of the tree on training data. The procedure produces trees that are invariant under any order preserving transformation of the predictor attributes. The CART mechanism is intended to produce not one, but a sequence of nested pruned trees, all of which are candidate optimal trees. The "right sized" or "honest" tree is identified by evaluating the predictive performance of every tree in the pruning sequence. CART offers no internal performance measures for tree selection based on the training data as such measures are deemed suspect. Instead, tree performance is always measured on independent test data and tree selection proceeds only after test-data-based evaluation. If no test data exist and cross validation has not been performed, CART will remain agnostic regarding which tree in the sequence is best. This is in sharp contrast to methods such as C4.5 that generate preferred models on the basis of training data measures.

III. RELATED WORKS

Although, using data mining in higher education is a recent research field, there are many works in this area. That is because of its potentials to educational institutes. Educational data mining is a promising area of research and it has a specific requirements not presented in other domains. Thus, work should be oriented towards educational domain of data mining. The goal of this study is to show how useful data mining can be used in higher education to improve student' performance.

Romero and Ventura [6], have a survey on educational data mining between 1995 and 2005. They concluded that educational data mining is a promising area of research and it has a specific requirements not presented in other domains. Thus, work should be oriented towards educational domain of data mining.

El-Halees [5], gave a case study that used educational data mining to analyze students' learning behavior. The goal of his study is to show how useful data mining can be used in higher education to improve student' performance. He used students' data from database course and collected all available data including personal records and academic records of students, course records and data came from e-learning system. Then, he applied data mining techniques to discover many kinds of knowledge such as association rules and classification rules using decision tree. Also he clustered the student into groups using EMclustering, and detected all outliers in the data using outlier analysis. Finally, he presented how can we benefited from the discovered knowledge to improve the performance of student.

Al-Radaideh et al. [1], applied the data mining techniques, particularly classification to help in improving the quality of the higher educational system by evaluating student data to study the main attributes that may affect the student performance in courses. The extracted classification rules are based on the decision tree as a classification method; the extracted classification rules are studied and evaluated. It allows students to predict the final grade in a course under study.

Baradwaj and Pal [3], applied the classification as data mining technique to evaluate student' performance, they used decision tree method for classification. The goal of their study is to extract knowledge that describes students' performance in end semester examination. They used students' data from the student' previous database including Attendance, Class test, Seminar and Assignment marks. This study helps earlier in identifying the dropouts and students who need special attention and allow the teacher to provide appropriate advising.

Shannaq et al. [7], applied the classification as data mining technique to predict the numbers of enrolled students by evaluating academic data from enrolled students to study the main attributes that may affect the students' loyalty (number of enrolled students). The extracted classification rules are based on the decision tree as a classification method, the extracted classification rules are studied and evaluated using different evaluation methods. It allows the University management to prepare necessary resources for the new enrolled students and indicates at an early stage which type of students will potentially be enrolled and what areas to concentrate upon in higher education systems for support.

Chandra and Nandhini [4], applied the association rule mining analysis based on students' failed courses to identifies students' failure patterns. The goal of their study is to identify hidden relationship between the failed courses and suggests relevant causes of the failure to improve the low capacity students' performances. The extracted association rules reveal some hidden patterns of students' failed courses which could serve as a foundation stone for academic planners in making academic decisions and an aid in the curriculum re-structuring and modification with a view to improving students' performance and reducing failure rate.

Ayesha et al. [2], used k-means clustering algorithm as a data mining technique to predict students' learning activities in a students' database including class quizzes, mid and final exam and assignments. These correlated information will be conveyed to the class teacher before the conduction of final exam. This study helps the teachers to reduce the failing ratio by taking appropriate steps at right time and improve the performance of students.

IV. DATA MINING PROCESS

Now-a-days, performance of individual student in educational system is evaluated based on the internal assessment and university examination. The internal assessment is calculated based on the performance of student in educational activities such as internals, assignment, seminars presented, entire lab work, counseling, technology acceptance and parental interaction. The internal assessment is calculated by teachers. The university examination is one that is scored by the student in semester examination. To pass university examination, each and every student has to gain minimum marks both in internal as well as final examination in semester.

A. Data preparations

The data set used in this study was obtained from B.TECH students of KITS engineering college from session 2006 to 2010. Initially size of the data is 30. In this step data stored in different tables was joined in a single table after joining process errors were removed.

Variable	Description	Possible Values
ESM	Earlier Semester Marks	{First > 60% Second >50 & <60% Third >40 & <50% Fail < 40%}
SPO	Seminar Performance outside the campus	{Poor , Average, Good}
ATT	Attendance	{Poor , Average, Good}
ELW	Entire Lab Work(internal + external)	{Yes, No}
INTERNALS	Average of Mid exams conducted internally	{Poor , Average, Good}
ASS	Assignment	{Yes, No}
COU	Counseling	{Yes, No}
TA	Technology acceptance	{Yes, No}
PI	Parental Interaction	{Yes, No}
UE	University examination	{First > 60% Second >50 & <60% Third >40 & <50% Fail < 40%}

The domain values for some of the variables were defined for the present investigation as follows:

ESM: ESM means earlier semester marks in B.Tech course. It is split into four class values: First > 60% , Second >50 & <60% , Third >40 & <50% and Fail < 40%.

INTERNALS: Internals means average of mid exams conducted internally in college. In each semester, two written exams and two online exams are conducted and averages of four tests are used to calculate internal marks. It is split into three class values: Poor, Average and Good.

SPO: Seminar performance participated outside the campus. Seminar performance is evaluated into three classes: *Poor – Presentation and communication skill is low, Average – Either presentation is fine or Communication skill is fines, Good – Both presentation and Communication skill is fine.*

ATT: Attendance of student. To attend the university exams, each and every student must have minimum 75% attendance. Attendance is divided into three classes: poor < 60%, average 60% > & < 75%and good >75%.

ASS – Assignment performance. Assignment performance is divided into two classes: *Yes – student submitted assignment, No – Student not submitted assignment.*

ELW – Entire Lab Work. Entire lab work means both internal and external lab work and exams. Entire Lab work is divided into two classes: *Yes – student completed lab work, No – student not completed lab work.*

COU- Counseling conducted to student in college. It is split into two class values: yes- Student attended to counseling, No- Student does not attended to counseling.

TA - Technology acceptance by both students and parents. It is divided into two classes: *Yes- Accepted, No- Doesn't accept.*

PI - Parental interaction. Parental interaction means whether parents are interested in knowing student status in college. It is dividing into two classes: yes- Interest in interacting, No- Doesn't interest in interacting.

UE – University exams obtained in B.Tech and it is declared as response variable. It is split into four class values: First > 60% , Second >50 & <60% , Third >40 & <50% and Fail < 40%.

B. Basic Framework

The new approach defines a new formal environment to prune and group discovered associations integrating knowledge into specific mining process of association rules. It is composed of three main parts (as shown in Figure 1). Firstly, a basic mining process is applied over data extracting a set of association rules. Secondly, the knowledge base allows formalizing user knowledge and goals. Domain knowledge allows a general view over user knowledge in database domain, and user expectations express user already knowledge over the discovered rules. Finally, the post-processing step consists in applying several operators (i.e. pruning) over user expectations in order to extract the interesting rules.

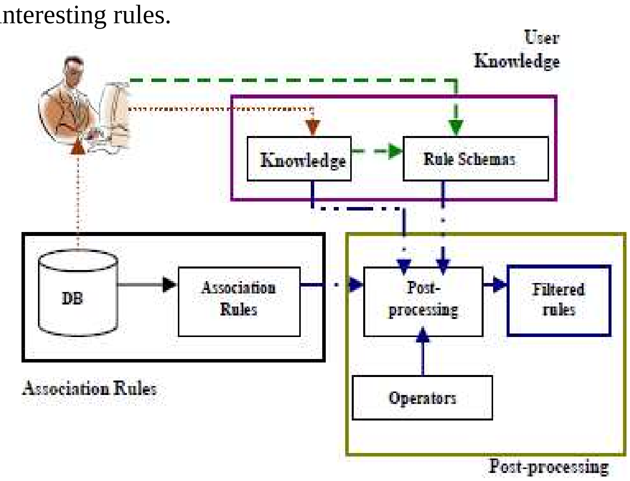

Fig. 1: Framework description

The novelty of this approach resides in supervising the knowledge discovery process using two different conceptual structures for user knowledge representation: integrate user knowledge in the post processing task and several rule schemas generalizing general impressions, and proposing an iterative process.

C. Interactive post mining process

The framework proposes to the user an interactive process of rule discovery. Taking into account his/her feedbacks, the user is able to revise his/her expectations in function of intermediate results. Several steps are suggested to the user in the framework as follows:

1. Knowledge construction—starting from the database, and eventually, from existing Knowledge, the user develops knowledge on database items;
2. Defining Rule Schemas (as GIs and RPCs)—the user expresses his/her local goals and expectations concerning the association rules that he/she wants to find;
3. Choosing the right operators to be applied over the rule schemas created, and then, applying the operators;
4. Visualizing the results—the filtered association rules are proposed to the user;
5. Selection/validation—starting from these preliminary results, the user can validate the results or he/she can revise his/her information;
6. Filters can be applied over rules whenever the user needs them with the main goal of reducing the number of rules; and
7. The interactive loop permits to the user to revise the information that he/she proposed. Thus, he/she can return to step 2 in order to modify the rule schemas, or he/she can return to step 3 in order to change the operators. Moreover, in the interactive loop, the user could decide to apply one of the two predefined filters discussed in step 6.

To improve association rule selection, we propose a new rule filtering model, called Rule Schemas (RS). A rule schema allows user expectations representation and permits to the user to supervise association rule mining, meanwhile operators guide the post-processing task by pruning and filtering discovered rules. The Rule Schema formalism is based on the specification language for user knowledge.

D. Operations over Rule Schemas

The rule schema filter is based on operators applied over rule schemas allowing the user to perform several actions over the discovered rules. We propose two important operators: pruning and filtering operators. The filtering operator is composed of three different operators: conforming, unexpectedness, and exception.

Pruning. The pruning operator allows to the user to remove families of rules that he/she considers uninteresting. In databases, there exist, in most cases, relations between items that we consider obvious or that we already know. Thus, it is not useful to find these relations among the discovered associations. The pruning operator applied over a rule schema, P(RS), eliminates all association rules matching the rule schema. To extract all the rules matching a rule schema, the conforming operator is used.

Conforming. The conforming operator applied over a rule schema, C(RS), confirms an implication or finds the implication between several concepts. As a result, rules matching all the elements of a nonimplicative rule schema are filtered. For an implicative rule schema, the condition and the conclusion of the association rule should match those of the schema.

Unexpectedness. With a higher interest for the user, the unexpectedness operator U(RS) proposes to filter a set of rules with a surprise effect for the user. This type of rules interests the user more than the conforming one since, generally, a decision-maker searches to discover new knowledge with regard to his/her prior knowledge.

Exceptions. Finally, the exception operator is defined only over implicative rule schemas (i.e., RS1) and extracts conforming rules with respect to the following new implicative rule schema: $X \wedge Z \rightarrow \neg Y$, where Z is a set of items.

In order to reduce the number of rules, three filters integrate the framework: operators applied over rule schemas, minimum improvement constraint filter, and item-relatedness filter. Minimum improvement constraint filter (MICF) selects only those rules whose confidence is greater with minimum than the confidence of any of its simplification s. The item-relatedness filter (IRF) Starting from the idea that the discovered rules are generally obvious, they introduced the idea of relatedness between items measuring their semantic distance in item taxonomies. This measure computes the relatedness of all the couples of rule items. We can notice that we can compute the relatedness for

IJSCE

the items of the condition or/and the consequent, or between the condition and the consequent of the rule.

S.N	ESM	INT	ATT	ASS	SPO	ELW	COU	TA	PI	UE
1	I	G	G	Y	Y	Y	Y	Y	Y	I
2.	I	G	A	Y	N	Y	Y	N	N	I
3.	I	G	A	N	N	N	N	Y	N	I
4.	I	A	G	N	N	Y	Y	N	N	I
5.	I	A	A	N	Y	N	Y	N	N	I
6.	I	A	A	N	N	Y	Y	N	Y	I
7	I	P	G	Y	Y	Y	Y	N	N	II
8	I	P	A	Y	Y	N	Y	N	N	II
9	I	A	A	N	Y	N	N	Y	N	II
10	II	G	G	Y	Y	Y	Y	Y	Y	I
11	II	G	G	Y	N	Y	Y	N	N	I
12	II	A	G	Y	Y	Y	N	N	N	II
13	II	G	A	Y	N	N	Y	N	N	II
14	II	A	A	N	N	N	Y	N	N	III
15	II	A	P	N	N	Y	Y	N	N	III
16	II	P	P	N	N	N	N	N	N	F
17	III	G	G	Y	Y	Y	Y	Y	Y	I
18	III	G	A	N	N	Y	Y	N	Y	II
19	III	A	A	Y	N	Y	N	N	N	II
20	III	A	A	Y	N	N	N	N	N	III
21	III	P	P	N	N	N	N	N	N	F
22	F	G	G	Y	Y	Y	Y	Y	Y	I
23	F	A	G	Y	N	Y	Y	N	N	II
24	F	A	A	Y	N	Y	N	N	N	II
25	F	A	A	N	N	N	Y	N	N	III
26	F	A	G	Y	N	N	Y	N	N	III
27	F	P	P	N	N	N	N	N	N	F
28	I	P	P	N	N	N	N	N	N	F
29	II	P	A	N	Y	N	N	N	N	F
30	III	P	A	N	Y	N	Y	N	N	F

‘G’: Good, ‘A’: Average, ’P’:Poor, ’Y’:Yes, ’N’:No, ’F’:Fail, ‘I’: First,’II’:Second,’III’:Third

V. RESULTS AND DISCUSSION

The data set of 30 students used in this study was obtained from B.Tech. A rule schema allows user expectations representation and permits to the user to supervise association rule mining, meanwhile operators guide the post-processing task by pruning and filtering discovered rules. The Rule Schema formalism is based on the specification language for user knowledge. From the above data set some rule schemas are listed below:

- IF ESM = “First‟ AND ATT = “Good‟ AND INT = “Good‟ AND ELW=”yes” AND ASS=”yes” AND COU=”yes” AND TA=”YES” THEN UE = “First”
- IF ESM = “First‟ AND ATT = “Average‟ AND INT = “Good‟ AND ELW=”no” AND ASS=”yes” AND COU=”yes” AND TA=”no” THEN UE = “First”
- IF ESM = “First‟ AND ATT = “Average‟ AND INT = “average‟ AND ELW=”no” AND ASS=”no” COU=”no” AND TA=” yes” THEN UE = “First”
- IF ESM = “First‟ AND ATT = “Poor‟ AND INT = “average‟ AND ELW=”no” AND ASS=”yes” AND COU=”no” AND TA=”no” THEN UE = “Second”
- IF ESM = “Second‟ AND ATT = “average‟ AND INT = “average‟ AND ELW=”yes” AND ASS=”no” AND COU=”yes” AND TA=”no” THEN UE = “Second”
- IF ESM = “Second‟ AND ATT = “average‟ AND INT = “average‟ AND ELW=”no” AND ASS=”yes” AND COU=”no” AND TA=”yes” THEN UE = “Second”
- IF ESM = “Second‟ AND ATT = “Good‟ AND INT = “average‟ AND ELW=”yes” AND ASS=”yes” AND COU=”no” AND TA=”no” THEN UE = “Third”
- IF ESM=”Third” AND ATT=”Good” AND INT=”Good” AND ELW=”yes” AND ASS=”yes” AND COU=”yes” AND TA=”yes” THEN UE=”First”
- IF ESM=”Third” AND ATT=”average” AND INT=”average” AND ELW=”yes” AND ASS=”yes” AND COU=”yes” AND TA=”yes” THEN UE=”second”
- IF ESM=”Third” AND ATT=”average” AND INT=”average” AND ELW=”yes” AND ASS=”yes” AND COU=”no” AND TA=”no” THEN UE=”Third”
- IF ESM=”Third” AND ATT=”Poor” AND INT=”Poor” AND ELW=”yes” AND ASS=”yes” AND COU=”no” AND TA=”no” THEN UE=”Fail”
- IF ESM=”fail” AND ATT=”Good” AND INT=”Good” AND ELW=”yes” AND ASS=”yes” AND COU=”yes” AND TA=”yes” THEN UE=”First”
- IF ESM=”Fail” AND ATT=”average” AND INT=”average” AND ELW=”yes” AND ASS=”yes” AND COU=”yes” AND TA=”yes” THEN UE=”second”
- IF ESM=”Fail” AND ATT=”Poor” AND INT=”Poor” AND ELW=”yes” AND ASS=”yes” AND COU=”no” AND TA=”no” THEN UE=”Fail”

The rule schema filter is used to retain the minimal rule schemas. Filter operators applied over rule schemas allowing the user to perform several actions over the discovered rules. We propose two important operators: pruning and filtering operators. While the pruning operator, we will remove families of rules which are uninteresting. After pruning, Minimum improvement constraint filter selects only those rules whose confidence is greater with minimum than the confidence of any of its simplification. The item-relatedness filter (IRF) finds relatedness between items measuring their semantic distance in item taxonomies. After applying the filters, finally generated rule schema is:

- IF ESM=”First/Second” AND INT=”Good” AND ATT=”Good/Average” AND ASS=”Yes” AND COU=”YES” THEN UE=”first”
- IF ESM=”First/Second” AND INT=”Average” AND ATT=”Average” AND ASS=”Yes” AND COU=”yes” THEN UE=”Second”

- IF ESM=”Second/Third” AND INT=”Average” AND ATT=”Good” AND ASS=”No” AND TA=”Yes” THEN UE=”Second”
- IF ESM=”Third” AND INT=”Average” AND ATT=”Average” AND ASS=”No” AND COU=”Yes” THEN UE=”Third”
- IF ESM=”Fail” AND INT=”Good” AND ATT=”Good” AND ASS=”Yes” AND COU=”No/Yes” THEN UE=” Second”
- IF ESM=”First/second/Third/Fail” AND INT=”Poor” AND ATT=”Poor” AND ASS=”No” THEN UE=”Fail”
- IF COU=”Yes” AND TA=”Yes” THEN PI=”Yes”
- IF COU=”No” AND TA=”No” THEN PI=”No”
- IF ESM=”Fail” AND PI=”Yes” AND INT=”Good” AND ASS=”Yes” THEN UE=”First/Second/Third”

VI. CONCLUSION

Data mining task is used in computer science and information technology aspects such as online learning and collaborative learning to facilitate students learning. In this paper, we use Association Rules Instead of tree based classification. A result of a tree based classification is complicated to understand and depends on the technical competency of the decision maker. Among sets of items in transaction databases, Association Rules aims at discovering implicative tendencies that can be valuable information for the decision-maker which is absent in tree based classifications. So we propose a new interactive approach to prune and filter discovered rules. First, we propose to integrate user knowledge in the post processing task. Second, we propose a Rule Schema formalism extending the specifications to obtain association rules from knowledge base. Compared to tree based classifications the results are better to understand and can be applied to real time use.

REFERENCES

[1] Al-Radaideh, Q., Al-Shawakfa, E. and Al-Najjar, M. (2006) ‘Mining Student Data Using Decision Trees’, The 2006 International Arab Conference on Information Technology (ACIT'2006) – Conference Proceedings.

[2] Ayesha, S. , Mustafa, T. , Sattar, A. and Khan, I. (2010) ‘Data Mining Model for Higher Education System’, European Journal of Scientific Research, vol. 43, no. 1, pp. 24-29.

[3] Baradwaj, B. and Pal, S. (2011) ‘Mining Educational Data to Analyze Student s’ Performance’, International Journal of Advanced Computer Science and Applications, vol. 2, no. 6, pp. 63-69.

[4] Chandra, E. and Nandhini, K. (2010) ‘Knowledge Mining from Student Data’, European Journal of Scientific Research, vol. 47, no. 1, pp. 156-163.

[5] El-Halees, A. (2008) ‘Mining Students Data to Analyze Learning Behavior: A Case Study’, The 2008 international Arab Conference of Information Technology (ACIT2008) – Conference Proceedings, University of Sfax, Tunisia, Dec 15- 18.

[6] Romero, C. and Ventura, S. (2007) ‘Educational data Mining: A Survey from 1995 to 2005’, Expert Systems with Applications (33), pp. 135-146.

[7] Shannaq, B. , Rafael, Y. and Alexandro, V. (2010) ‘Student Relationship in Higher Education Using Data Mining Techniques’, Global Journal of Computer Science and Technology, vol. 10, no. 11, pp. 54-59.

[8] S. T. Hijazi, and R. S. M. M. Naqvi, “Factors affecting student□s performance: A Case of Private Colleges”, Bangladesh e-Journal of Sociology, Vol. 3, No. 1, 2006.

[9] U. K. Pandey, and S. Pal, “A Data mining view on class room teaching language”, (IJCSI) International Journal of Computer Science Issue, Vol. 8, Issue 2, pp. 277-282, ISSN:1694-0814, 2011.

[10] Shaeela Ayesha, Tasleem Mustafa, Ahsan Raza Sattar, M. Inayat Khan, “Data mining model for higher education system”, Europen Journal of Scientific Research, Vol.43, No.1, pp.24-29, 2010.

BIOGRAPHY

Mr. P. AJITH , [Research Scholar , KL University] received B.Sc. degree from Nagarjuna University, Guntur, India in 1993, Master of Computer Applications from Sardar Patel University, Gujarat, India in 1996. MTech(CSE) from Bharat University, Chennai, India in 2006 and is currently pursuing Ph.D. at KL University, Vijayawada, A.P., India.

He is currently working as Head & Associate Professor in the Department of Computer Science, in Sriji College, Ongole, A.P, India. His area of interests is Data Mining & its applications.

B. Tejaswi , received Bachelor of Technology degree in Computer Science & Engineering from Jawaharlal Nehru Technological University , Kakinada, Andhra Pradesh , India in 2010, Master of Technology in Computer Science & Engineering from Jawaharlal Nehru Technological University , Kakinada, Andhra Pradesh India in 2012. She is currently working Asst. Professor in the Department of Computer Science & Engineering, in KKR & KSR Institute of Technology & Sciences, Guntur, India.

Dr. M. Srinivasa Sesha Sai , received B.Sc. degree from Nagarjuna University, India in 1994, Master of Computer Applications from Acharya Nagarjuna University, India in 1997, Master of Philosophy in Computer Science from Alagappa University, India in 2004,Master of Technology in Computer Science & Engineering from Dr. M.G.R. University, India in 2007 and Ph.D. in Computer Science & Engineering from Acharya Nagarjuna University, India in 2011. He is currently working as Head & Professor in the Department of Computer Science & Engineering, in KKR & KSR Institute of Technology & Sciences, Guntur, India. His area of interests are Data Mining & its applications, Image Processing and Pattern Recognitions and Digital Communications