Data Mining Techniques in EDM for Predicting the Performance of Students

Ajay Kumar Pal
Research Scholar,
Sai Nath University,
Ranchi, Jharkhand, India

Saurabh Pal
Head, Dept. of MCA,
VBS Purvanchal University,
Jaunpur, UP, India
Email: drsaurabhpal {at} yahoo.co.in

Abstract— **In recent, growth of higher education has increased rapidly. Many new institutions, colleges and universities are being established by both the private and government sectors for the growth of education and welfare of the students. Each institution aims at producing higher and exemplary education rates by employing various teaching and grooming methods. But still there are cases of unemployment that exists among the medium and low risk students.**

This paper describes the use of data mining techniques to improve the efficiency of academic performance in the educational institutions. Various data mining techniques such as decision tree, association rule, nearest neighbors, neural networks, genetic algorithms, exploratory factor analysis and stepwise regression can be applied to the higher education process, which in turn helps to improve student's performance. This type of approach gives high confidence to students in their studies. This method helps to identify the students who need special advising or counseling by the teacher which gives high quality of education.

Keywords-component; Data Mining; KDD; EDM; Association Rule

I. INTRODUCTION

Data mining is the powerful technology for analyzing important information from the data warehouse. It is data analysis methodology used to identify hidden patterns in a large data set. Data mining is one of the steps in KDD process. Knowledge discovery (KDD) aims at the discovery of useful information from large collections of data [2]. The main goal of data mining in the KDD process concerned with the algorithmic means by which patterns or structures are enumerated from the data under acceptable computational efficiency limitations. Data mining has been successfully used in different areas including the educational environment. Educational data mining is an interesting research area which extracts useful, previously unknown patterns from educational database for better understanding, improved educational performance and assessment of the student learning process [3].

Data miming consists of a set of techniques that can be used to extract relevant and interesting knowledge from data. Data mining has several tasks such as association rule mining, classification and prediction, and clustering. Classification techniques are supervised learning techniques that classify data item into predefined class label. It is one of the most useful techniques in data mining to build classification models from an input data set. The used classification techniques commonly build models that are used to predict future data trends.

There are increasing research interests in education field using data mining. Application of Data mining techniques concerns to develop the methods that discover knowledge from data and used to uncover hidden or unknown information that is not apparent, but potentially useful [8]. The data can be personal or academic which can be used to understand students behavior to assist instructors, to improve teaching, to improve curriculums and many other benefits.

There are a large number of research papers on educational data mining discussing various problems within the higher education sector and providing examples for successful solutions was done by using data mining. C Romero and S Vetura [17] made a comprehensive study on the development of this educational data mining since 1995 to 2005. Their paper surveys the application of data mining to traditional education systems, particular web-based course, well known learning content management systems and adaptive and intelligent web-based educational systems. Mostly the problems attracting the attention of researchers are focused mainly on retention of students, more effective targeted marketing, improving institutional efficiency, and alumni management.

This study investigates and compares the educational domain of data mining from data that come from students personal, social, psychological and other environmental variables. The scope of this research paper, makes to extract the knowledge discover from the student database for improving the student performance. Here by, data mining techniques including a rule learner (OneR), a common decision tree algorithm C4.5 (J48), a neural network (MultiLayer Perceptron), and a Nearest Neighbour algorithm (IB1) are used.

II. BACKGROUND AND RELATED WORK

Due to high accuracy and prediction quality data mining technique is widely used in different areas. Education sector is also enriched with the help of this technique. A number of

journal and literature are available containing educational data mining. Few of them are listed below for reference.

Bharadwaj and Pal [1] conducted study on the student performance by selecting 300 students from 5 different degree colleges in India. In their study, it was found that students' grade in senior secondary exam, living location, medium of teaching, mother's qualification, family annual income, and student's family status were highly correlated with the student academic performance.

Bharadwaj and Pal [4] in their another study they used students' previous semester marks, class test grade, seminar performance, assignment performance, general proficiency, attendance in class and lab work to predict students' mark in their end semester.

Kovacic [5] used enrollment data to predict successful and unsuccessful student in New Zealand and he found 59.4% and 60.5% of classification accuracy while using decision tree algorithms CHAID and CART respectively.

Yadav, Bhardwaj and Pal [6] conducted study on the student retention based by selecting 398 students from MCA course of VBS Purvanchal University, Jaunpur, India. By means of classification they show that student's graduation stream and grade in graduation play important role in retention.

Pal [7] conducted study on the student dropout rate by selecting 1650 students from different branches of engineering college. In their study, it was found that student's dropout rate in engineering exam, high school grade; senior secondary exam grade, family annual income and mother's occupation were highly correlated with the student academic performance.

Yadav and Pal [9] conducted a study using classification tree to predict student academic performance using students' gender, admission type, previous schools marks, medium of teaching, location of living, accommodation type, father's qualification, mother's qualification, father's occupation, mother's occupation, family annual income and so on. In their study, they achieved around 62.22%, 62.22% and 67.77% overall prediction accuracy using ID3, CART and C4.5 decision tree algorithms respectively.

In another study Yadav *et al*. [10] used students' attendance, class test grade, seminar and assignment marks, lab works to predict students' performance at the end of the semester with the help of three decision tree algorithms ID3, CART and C4.5. In their study they achieved 52.08%, 56.25% and 45.83% classification accuracy respectively.

Merceron A et al. [11] concluded that association technique requires not only that adequate thresholds be chosen for the two standard parameters of support and confidence, but also that appropriate measures of interestingness be considered to retain meaning rules that filter uninterestingness ones out.

Oladipupo O.O. and Oyelade O.J. [12] study has bridge the gap in educational data analysis and shows the potential of the association rule mining algorithm for enhancing the effectiveness of academic planners and level advisers in higher institutions of learning.

Bray [13], in his study on private tutoring and its implications, observed that the percentage of students receiving private tutoring in India was relatively higher than in Malaysia, Singapore, Japan, China and Sri Lanka. It was also observed that there was an enhancement of academic performance with the intensity of private tutoring and this variation of intensity of private tutoring depends on the collective factor namely socio-economic conditions.

Galit [14] gave a case study that use students data to analyze their learning behavior to predict the results and to warn students at risk before their final exams.

Al-Radaideh, et al [15] applied a decision tree model to predict the final grade of students who studied the C++ course in Yarmouk University, Jordan in the year 2005. Three different classification methods namely ID3, C4.5, and the NaïveBayes were used. The outcome of their results indicated that Decision Tree model had better prediction than other models.

III. DATA MINING TECHNIQUES

This paper uses a rule learner (OneR), a common decision tree algorithm C4.5 (J48), a neural network (MultiLayer Perceptron), and a Nearest Neighbour algorithm (IB1). These classification algorithms are selected because they are very often used for research purposes and have potential to yield good results. Moreover, they use different approaches for generating the classification models, which increases the chances for finding a prediction model with high classification accuracy. The OneR Rule Learner algorithm produces a one-level decision tree expressed in the form of a set of rules that all test one particular attribute – the minimum classification-error attribute. It is a simple, cheap method that often produces good rules with high accuracy. The Decision Tree algorithms generate models in the form of a tree-like structure, which starts from root attributes and ends with leaf nodes, describing the relationship among attributes and the relative importance of attributes. They represent rules which could easily be understood and interpreted by users, do not require complex data preparation, and perform well for numerical and categorical variables. Neural networks produce classification models in the form of a mathematical model, consisting of interconnected computational elements (neurons) and processing information using a connectionist approach to computation. They are used to model complex relationships between inputs and outputs and very often yield very good results. The K-Nearest Neighbor algorithm (k-NN) is a method for classifying instances based on measuring the distance between the classified instance and the closest training examples in the feature space. It is easily understood by users, often provides good classification results and performs well for large datasets.

A. OneR (Rule Learner)

OneR, short for "One Rule", is a simple, yet accurate, classification algorithm that generates one rule for each predictor in the data, then selects the rule with the smallest total

error as its "one rule". To create a rule for a predictor, we construct a frequency table for each predictor against the target. It has been shown that OneR produces rules only slightly less accurate than state-of-the-art classification algorithms while producing rules that are simple for humans to interpret.

B. C4.5

This algorithm is a successor to ID3 developed by Quinlan Ross [16]. It is also based on Hunt's algorithm.C4.5 handles both categorical and continuous attributes to build a decision tree. In order to handle continuous attributes, C4.5 splits the attribute values into two partitions based on the selected threshold such that all the values above the threshold as one child and the remaining as another child. It also handles missing attribute values. C4.5 uses Gain Ratio as an attribute selection measure to build a decision tree. It removes the biasness of information gain when there are many outcome values of an attribute.

At first, calculate the gain ratio of each attribute. The root node will be the attribute whose gain ratio is maximum. C4.5 uses pessimistic pruning to remove unnecessary branches in the decision tree to improve the accuracy of classification.

C. MultiLayer Perceptron

Multilayer Perceptron (MLP) algorithm is one of the most widely used and popular neural networks. Multilayer Perceptron (MLP) is a feed forward artificial neural network model that maps sets of input data onto a set of appropriate output. An MLP consists of multiple layers of nodes in a directed graph, with each layer fully connected to the next one. Their current output depends only on the current input instance. It trains using back propagation.

D. Nearest Neighbour algorithm

IB1 is nearest neighbour classifier. It uses normalized Euclidean distance to find the training instance closest to the given test instance, and predicts the same class as this training instance. If several instances have the smallest distance to the test instance, the first one obtained is used. Nearest neighbour method is one of the effortless and uncomplicated learning/classification algorithms, and has been effectively applied to a broad range of problems.

IV. DATA MINING PROCESS

In this study, data gathered from different degree colleges and institutions affiliated with Dr. R. M. L. Awadh University, Faizabad, India. These data are analyzed using association rules and decision trees to predict the student's performance. In order to apply this technique following steps are performed in sequence:

A. Data Preparations

The data set used in this study was obtained from different colleges on the sampling method for B.Sc. (Bachelors of Science) course of session 2011-12. Initially size of the data is 200. In this step data stored in different tables was joined in a single table after joining process errors were removed.

B. Data selection and transformation

In this step only those fields were selected which were required for data mining. A few derived variables were selected. While some of the information for the variables was extracted from the database. All the predictor and response variables which were derived from the database are given in Table I for reference.

TABLE I. : STUDENT RELATED VARIABLES

Variables	Description	Possible Values
Sex	Students Sex	{Male, Female}
Cat	Students category	{Unreserved, OBC, SC, ST}
HSG	Students grade in High School	{O – 90% -100%, A – 80% - 89%, B – 70% - 79%, C – 60% - 69%, D – 50% - 59%, E – 40% - 49%, F - < 40%}
SSG	Students grade in Senior Secondary	{O – 90% -100%, A – 80% - 89%, B – 70% - 79%, C – 60% - 69%, D – 50% - 59%, E – 40% - 49%, F - < 40% }
Atype	Admission Type	{Test, Direct}
Med	Medium of Teaching	{Hindi, English}
CLoc	College Location	{ Village, Town, Tahseel, District}
PCol	Profile of College	{Good, Bad}
SelfC	Self Center	{Yes, No}
LLoc	Living Location of Student	{Village, Town, Tahseel, District}
Hos	Student live in hostel or not	{Yes, No}
FSize	student's family size	{1, 2, 3, >3}
FStat	Students family status	{Joint, Individual}
FAIn	Family annual income	{BPL, poor, medium, high}
FQual	Fathers qualification	{no-education, elementary, secondary, UG, PG, Ph.D. NA}
MQual	Mother's Qualification	{no-education, elementary, secondary, UG, PG, Ph.D. NA}
FOcc	Father's Occupation	{Service, Business, Agriculture, Retired,NA}
MOcc	Mother's Occupation	{House-wife (HW), Service, Retired, NA}
Result	Result in B.Sc.	{First ≥ 60% Second ≥ 45 & <60% Third ≥ 36 & < 45%, Fail< 36%}

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

- Cat – From ancient time Indians are divided in many categories. These factors play a direct and indirect role in the daily lives including the education of young people. Admission process in India also includes different percentage of seats reserved for different categories. In terms of social status, the Indian population is grouped into four categories: General, Other Backward Class (OBC), Scheduled Castes (SC) and Scheduled Tribes (ST). Possible values are Unreserved, OBC, SC and ST.

- HSG - Students grade in High School education. Students who are in state board appear for six subjects each carry 100 marks. Grade are assigned to all students using following mapping O – 90% to 100%, A – 80% - 89%, B – 70% - 79%, C – 60% - 69%, D – 50% - 59%, E – 40% - 49%, and F - < 40%.
- SSG - Students grade in Senior Secondary education. Students who are in state board appear for five subjects each carry 100 marks. Grade are assigned to all students using following mapping O – 90% to 100%, A – 80% - 89%, B – 70% - 79%, C – 60% - 69%, D – 50% - 59%, E – 40% - 49%, and F - < 40%.
- Atype - The admission type which may be through Entrance Examination (EE) or direct admission through college procedure.
- Med – This paper study covers only the colleges of Uttar Pradesh state of India. Here, medium of instructions are Hindi or English.
- FSize-. According to population statistics of India, the average number of children in a family is 3.1. Therefore, the possible range of values is from one to greater than three.
- Result – Students result in B.Sc. Result is declared as response variable. It is split into four classes First, Second, Third and Fail.

C. Implementation of Mining Model

Weka is open source software that implements a large collection of machine leaning algorithms and is widely used in data mining applications. From the above data, per.arff file was created. This file was loaded into WEKA explorer. The classify panel enables the user to apply classification and regression algorithms to the resulting dataset, to estimate the accuracy of the resulting predictive model, and to visualize erroneous predictions, or the model itself. The rule learner (OneR), decision tree algorithm C4.5 (J48), neural network (MultiLayer Perceptron) and Nearest Neighbour algorithm (IB1) were implemented in WEKA. Under the "Test options", the 10-fold cross-validation is selected as our evaluation approach. Since there is no separate evaluation data set, this is necessary to get a reasonable idea of accuracy of the generated model. These predictive models provide ways to predict whether a new student will perform or not.

D. Results and Discussion

After applying the pre-processing and preparation methods, we try to analyze the data visually and figure out the distribution of values. Figure 1 depicts the distribution of values of students.

Figure 1: Visualization of the Students Categorization

We have carried out some experiments in order to evaluate the performance and usefulness of different classification algorithms for predicting students' placement. The results of the experiments are shown in table II.

TABLE II: PERFORMANCE OF THE CLASSIFIERS

Evaluation Criteria	**Classifiers**			
	OneR	**J48**	**MLP**	**IB1**
Timing to build model (in Sec)	0	0	8.59	0
Correctly classified instances	130	142	157	164
Incorrectly classified instances	70	58	43	36
Accuracy (%)	65.00%	78.00%	78.50%	82.00%

The percentage of correctly classified instances is often called accuracy or sample accuracy of a model. So IB1 classifier has more accuracy than other classifiers.

Kappa statistic, mean absolute error and root mean squared error will be in numeric value only. We also show the relative absolute error and root relative squared error in percentage for references and evaluation. The results of the simulation are shown in Tables III.

TABLE III: TRAINING AND SIMULATION ERROR

Evaluation Criteria	**Classifiers**			
	OneR	**J48**	**MLP**	**IB1**
Kappa statistic	0.4492	0.5635	0.6807	0.7345
Mean absolute error (MAE)	0.175	0.1693	0.1128	0.09
Root mean squared error (RMSE)	0.4183	0.3433	0.3004	0.3
Relative absolute error (RAE)	51.72%	50.03%	33.34%	26.60%
Root relative squared error (RRSE)	101.82%	83.55%	73.11%	73.01%

Once Predictive model is created, it is necessary to check how accurate it is, The Accuracy of the predictive model is calculated based on the precision, recall values of classification matrix.

PRECISION is the fraction of retrieved instances that are relevant. It is calculated as total number of true positives divided by total number of true positives + total number of false positives.

$$\text{Precision} = \frac{\text{True positives}}{\text{True positives} + \text{False positives}}$$

RECALL is fraction of relevant instances that are retrieved. It is usually expressed as a percentage. It is calculated as total number of true positives divided by total number of true positives + total number of false negatives.

$$\text{Recall} = \frac{\text{True positives}}{\text{True positives} + \text{False negatives}}$$

Comparison of evaluation measures by class are shown in table IV.

TABLE IV: COMPARISON OF EVALUATION MEASURES

Classifier	TP	FP	Precision	Recall	Class
OneR	0.904	0.282	0.692	0.892	First
	0.789	0.233	0.589	0.606	Second
	0.846	0.04	0.65	0.5	Fail
	0.55	0	0	0	Third
J48	0.807	0.137	0.807	0.807	First
	0.69	0.233	0.62	0.69	Second
	0.846	0.034	0.786	0.846	Fail
	0.2	0.033	0.4	0.2	Third
MLP	0.843	0.085	0.875	0.843	First
	0.775	0.163	0.724	0.775	Second
	0.846	0.017	0.88	0.846	Fail
	0.5	0.05	0.526	0.5	Third
IB1	0.904	0.06	0.915	0.904	First
	0.789	0.109	0.8	0.789	Second
	0.846	0.017	0.88	0.846	Fail
	0.55	0.067	0.478	0.55	Third

The performance of the learning techniques is highly dependent on the nature of the training data. Confusion matrices are very useful for evaluating classifiers. The columns represent the predictions, and the rows represent the actual class. To evaluate the robustness of classifier, the usual methodology is to perform cross validation on the classifier.

TABLE V: CONFUSION MATRIX

Classifier	First	Second	Fail	Third	Class
OneR	74	9	0	0	First
	24	43	4	0	Second
	2	11	13	0	Fail
	7	10	3	0	Third
J48	67	15	1	0	First
	15	49	2	5	Second
	0	3	22	1	Fail
	1	12	3	4	Third
MLP	70	12	0	1	First
	8	55	1	7	Second
	1	2	22	1	Fail
	1	7	2	10	Third
IB1	75	7	0	1	First
	6	56	1	8	Second
	9	1	22	3	Fail
	1	6	2	11	Third

Figures 2 and 3 are the graphical representations of the simulation result.

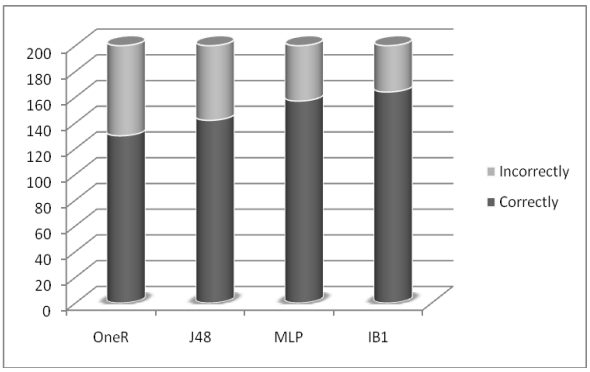

Figure 2: Efficiency of different models

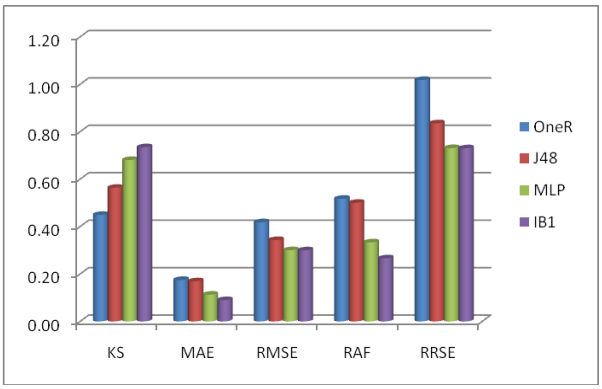

Figure 3: Comparison between Parameters

Based on the above Figures 2, 3 and Table II, we can clearly see that the highest accuracy is 82.00% and the lowest is 65.00%. The other algorithm yields an average accuracy of 78%. In fact, the highest accuracy belongs to the IB1 Classifier followed by Multilayer Perceptron function with a percentage of 78.50% and subsequently J48 tree. An average of 148 instances out of total 200 instances is found to be correctly classified with highest score of 164 instances compared to 130 instances, which is the lowest score. The total time required to build the model is also a crucial parameter in comparing the classification algorithm.

In this simple experiment, from Table II, we can say that a OneR, J48 and IB1 requires the shortest time which is around 0 seconds compared to MLP which requires the longest model building time which is around 8.59 seconds.

Kappa statistic is used to assess the accuracy of any particular measuring cases, it is usual to distinguish between the reliability of the data collected and their validity [18].

The average Kappa score from the selected algorithm is around 0.4-0.7. Based on the Kappa Statistic criteria, the accuracy of this classification purposes is substantial [18]. From Figure 3, we can observe the differences of errors resultant from the training of the three selected algorithms. This experiment implies a very commonly used indicator which is mean of absolute errors and root mean squared errors. Alternatively, the relative errors are also used. Since, we have two readings on the errors, taking the average value will be wise.

Decision trees are considered easily understood models because a reasoning process can be given for each conclusion. Knowledge models under this paradigm can be directly transformed into a set of IF-THEN rules that are one of the most popular forms of knowledge representation, due to their simplicity and comprehensibility which professor can easy understand and interpret. We can summarize the following results:

```
J48 pruned tree
------------------
SSG = A
|   FAIn = Medium
|   |   Atype = EE: First (33.0/4.0)
|   |   Atype = Direct
|   |   |   Sex = M: First (2.0)
|   |   |   Sex = F: Second (4.0/1.0)
|   FAIn = High: First (12.0)
|   FAIn = Poor: Second (2.0)
|   FAIn = BPL: First (0.0)
SSG = B
|   Mqual = UG
|   |   Cat = Unreserved: Second (12.0/5.0)
|   |   Cat = OBC: Second (0.0)
|   |   Cat = SC: First (1.0)
|   |   Cat = ST: Third (2.0)
|   Mqual = PG
|   |   Sex = M: First (8.0)
|   |   Sex = F: Second (3.0/2.0)
|   Mqual = Elementry
|   |   Hos = Yes: Fail (2.0/1.0)
|   |   Hos = No: Second (4.0)
|   Mqual = Secondry
|   |   FAIn = Medium: First (6.0)
|   |   FAIn = High: First (0.0)
|   |   FAIn = Poor: Second (2.0)
|   |   FAIn = BPL: First (0.0)
|   Mqual = Ph.D: First (2.0)
SSG = C
|   Fqual = PG: Second (4.0/1.0)
|   Fqual = UG
|   |   Sex = M
|   |   |   Pcol = Good: First (2.0)
|   |   |   Pcol = Bad: Second (14.0/5.0)
|   |   Sex = F: Second (12.0)
|   Fqual = Ph.D.: Second (2.0)
|   Fqual = Secondry: Fail (10.0/1.0)
|   Fqual = Elementry
|   |   Sex = M: Second (2.0)
|   |   Sex = F: First (2.0)
SSG = D
|   HSG = A: Second (13.0/2.0)
|   HSG = O: Second (0.0)
|   HSG = B: Second (0.0)
|   HSG = C: Fail (2.0)
|   HSG = E: Second (0.0)
|   HSG = D
|   |   Pcol = Good
|   |   |   Cloc = Distric: First (2.0/1.0)
|   |   |   Cloc = Village: Second (2.0)
|   |   |   Cloc = Tahseel: Second (0.0)
|   |   Pcol = Bad: Third (4.0)
|   HSG = F: Second (0.0)
SSG = O: First (12.0/1.0)
SSG = E
|   Hos = Yes: Fail (11.0)
|   Hos = No
|   |   Cat = Unreserved: Second (0.0)
|   |   Cat = OBC: Second (0.0)
|   |   Cat = SC
|   |   |   Atype = EE: Fail (2.0)
|   |   |   Atype = Direct: Third (2.0)
|   |   Cat = ST: Second (5.0/1.0)
SSG = F: Second (2.0)
```

V. CONCLUSIONS

Frequently used classifiers are studied and the experiments are conducted to find the best classifier for predicting the student's performance.

As a conclusion, we have met our objective which is to evaluate the performance of student by the four selected classification algorithms based on Weka. The best algorithm based on the placement data is IB1 Classification with an accuracy of 82.00% and the total time taken to build the model is at 0 seconds. IB1 classifier has the lowest average error at

0.20 compared to others. These results suggest that among the machine learning algorithm tested, IB1 classifier has the potential to significantly improve the conventional classification methods for use in performance.

Result shows that SSG (Students grade in senior secondary), HSG (Students grade in High School), Mqual (Mother's qualification) and FAIn (Family Annual Income) more affect the performance of the students.

The empirical results show that we can produce short but accurate prediction list for the student by applying the predictive models to the records of incoming new students. This study will also work to identify those students which needed special attention.

REFERENCES

[1] B.K. Bharadwaj and S. Pal. "Data Mining: A prediction for performance improvement using classification", International Journal of Computer Science and Information Security (IJCSIS), Vol. 9, No. 4, pp. 136-140, 2011.

[2] Heikki, Mannila, .Data mining: machine learning, statistics, and databases., IEEE, 1996.

[3] Moucary et,al.,Improving student performance using data clustering and neural networks in foreign language based higher education,The Research Bulletin of Jordan ACM, vol II (III).

[4] B.K. Bharadwaj and S. Pal. "Mining Educational Data to Analyze Students' Performance", International Journal of Advance Computer Science and Applications (IJACSA), Vol. 2, No. 6, pp. 63-69, 2011.

[5] Z. J. Kovacic, "Early prediction of student success: Mining student enrollment data", Proceedings of Informing Science & IT Education Conference 2010.

[6] S. K. Yadav, B.K. Bharadwaj and S. Pal, "Mining Educational Data to Predict Student's Retention :A Comparative Study", International Journal of Computer Science and Information Security (IJCSIS), Vol. 10, No. 2, 2012.

[7] Pal S., "Mining Educational Data to Reduce Dropout Rates of Engineering Students", I.J. Information Engineering and Electronic Business (IJIEEB), Vol. 4, No. 2, 2012, pp. 1-7.

[8] Pavel Berkhin, Survey of Clustering Data Mining Techniques, Accrue Software, Inc.

[9] S. K. Yadev & Pal., S. 2012. Data Mining: A Prediction for Performance Improvement of Engineering Students using Classification, World of Computer Science and Information Technology (WCSIT), 2(2), 51-56.

[10] S. K. Yadav, B. K. Bharadwaj & Pal, S. 2011. Data Mining Applications: A comparative study for predicting students' performance, International journal of Innovative Technology and Creative Engineering (IJITCE), 1(12).

[11] Merceron A, Yacef K, "Revisiting interestingness of strong symmetric association rules in educational data",Proceedings of the International Workshop on Applying Data Mining in e-Learning 2007.

[12] Oladipupo O.O., Oyelade O.J., "Knowledge Discovery from students' Result Repository: Association Rule Mining Approach", IJCSS Vol. 4: issue 2.

[13] M. Bray, .The shadow education system: private tutoring and its implications for planners., (2nd ed.), UNESCO, PARIS, France, 2007.

[14] Galit.et.al, "Examining online learning processes based on log files analysis: a case study". Research, Reflection and Innovations in Integrating ICT in Education 2007.

[15] Q. A. AI-Radaideh, E. W. AI-Shawakfa, and M. I. AI-Najjar, "Mining student data using decision trees", International Arab Conference on Information Technology(ACIT'2006), Yarmouk University, Jordan, 2006.

[16] Quinlan, J.R. (1993), C4.5: Programs for machine learning, Morgan Kaufmann, San Francisco.

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

[18] Kappa at http://www.dmi.columbia.edu/homepages/chuangj/ kappa.

Ajay Kumar Pal received his MCA. (Master of Computer Applications) from VBS Purvanchal University, Jaunpur, UP, India. Currently he is working as Head Of Department of Computer Application, Shri Vishwanath P.G. College Kalan Sultanpur(U.P.) At present, he is doing research in Data Mining and Knowledge Discovery. He is an active member of CSI and National Science Congress.

Saurabh Pal received his M.Sc. (Computer Science) from Allahabad University, UP, India (1996) and obtained his Ph.D. degree from the Dr. R. M. L. Awadh University, Faizabad (2002). He then joined the Dept. of Computer Applications, VBS Purvanchal University, Jaunpur as Lecturer. At present, he is working as Head and Sr. Lecturer at Department of Computer Applications.
Saurabh Pal has authored more than 35 numbers of research papers in international/national Conference/journals and also guides research scholars in Computer Science/Applications. He is an active member of CSI, Society of Statistics and Computer Applications and working as reviewer and member of editorial board for more than 15 international journals. His research interests include Image Processing, Data Mining, Grid Computing and Artificial Intelligence.