International Journal of Advances in Computer Science & Its Applications
Volume 6 : Issue 1 [ISSN 2250-3765]
Publication Date : 18 April, 2016

Students' Academic Counseling from Attribute Precedence Relations using EDM

[Mamta Singh[1], Jyoti Singh[2] , Arpana Rawal[3]]

***Abstract*— Educational Data Mining (EDM) techniques play an important role in understanding hidden students' data patterns to improve the quality of teaching-learning professions. In machine learning, feature selection usually emerges as a preprocessing step to extract necessary and sufficiently small subset of features for predictive / decision-making type of learning tasks. In this study, authors decided to work only upon external (changeable) attributes of students by assigning weights that reflect their academic efforts put in for those attributes. The attribute precedence levels extracted student-wise by current FE model due to academic efforts put up by students in their on-going course were compared with equivalently generated precedence relations from RELIEF method and it's variant. The favorable model accuracies of these precedence relations when compared with RELIEF have given a new meaning to EDM objectives in the direction of individual student counseling encouraging them to appraise themselves amidst their course tenure in right direction.**

Keywords*— EDM, Students' Academic Performance, Attribute Precedence Relations, Near-Hit, Near-Miss, RELIEF weights.

I. Introduction

Machine learning techniques in higher education are used to help to universities, colleges, instructors and the students getting better in their performance, consequently enveloped in term called Educational Data Mining (EDM). Assessment of Students' Academic Performance in various dimensions of Academia has emerged as a prime EDM objective for understanding students' responsive trends towards their courses. Owing to the sophisticated data management tools and revolution in IT, rapidly increasing levels of Educational databases, are provided by most of the Academic Institutions. Recent studies have attempted to find valuable information hidden in these students' databases using a series of statistical analyses and machine-learning techniques in data mining.

Mamta Singh
Department of Computer Science, Sai Mahavidyalaya
Bhilai, Chhattisgarh
India

Jyoti Singh
Directorate of Technical Education
Raipur, Chhattisgarh
India

Arpana Rawal
Department of Computer Science and Engineering, Bhilai Institute of Technology, Durg, Chhattisgarh
India

Although EDM has seen efforts of building offline Students' Recommender Systems and web-based decision-making tools, still the research realm suffers from partial isolation as most of the academic institutions are unaware or not able to realize the importance of using such tools to enhance students' learning levels. The work has already begun in the direction of predicting the learners' and educators' academic appraisals instead of constraining the mining objectives to admissions and enrollment procedures, detection of drop-outs and at-risk students. In this context, counseling the student on individual basis based upon his / her predicted performance in an academic program, before the commencement of final examinations is a difficult but poses to be a potential undertaking.

II. Overview of EDM models

Focusing the present work with reference to Higher Educational Universities, following case studies from various academic environments were reported to have been conducted by several mining researchers upon students' academic performance in higher technical education:

A. *Modeling type*

Pandey and Pal Investigated whether new comer students who wish to seek admission in PGDCA (one-year technical course run nation-wide for graduate students) will be performers or non-performers, a dichotomous classification model [1]. They were able to correlate the actual students' class attendance with 'Students' interestingness in hearing the Language medium of Class-room teaching' attribute using association rule mining methodology [2]. Yet other classification models aimed at resolving students' enrollment management system for seeking admissions in MCA (a technical course), predicting the drop-out status of students using DT classifiers as well as identifying, "Student's Retention for the pursued course" by predicting upon decision-tree classifiers: ID3, C4.5 and ADT [3] [4] [5]. To enrich the EDM modeling parameters, efforts were made to develop on-line educational assessment portals that assessed the student performances based on the homework assignments, quizzes and online- examinations, (Ex- LON-CAPA (Learning Online Network with Computer-Assisted Personalized Approach) developed by B. M. Bidgoli [6].

B. *Mining Method*

Early prediction modeling based on prominent parameters like internal test marks, attendance and influence of friendship were used to classify student instances as performers / under-performers with DT C5.0 classifier [7]. The works related to identifying correlation between predictor and response variables usually used ARM classification methods and invariably regression like machine-learning techniques [8]. Performance comparisons of classification tasks upon Students' academic performance revealed that at times Naïve

Bayesian classifier outperforms other k-NN and DT classifiers [9] [10].

C. Input-Output Parameters

Any kind of EDM framework begins with setting of Input Parameters that reflect the students' performance directly / indirectly from the instant, the students take up their courses. Normally, EDM predictor attributes range among students' personal profiles, demographic data, academic data and behavioral attributes: **Demographic**: Parents' Academic Qualification, Parents occupation, Family size, Family's annual income, Nature of accommodation. **Academic**: Grade in high school, Grade in higher secondary, Computer literacy, Exposure to programming, Stream of education, Previous Semester Marks, Class Test Grade, Seminar Attendance, Assignment, Lab Work, End Semester Marks , No. of backlogs. **Personal**: Gender, Age, Cast Category, Medium of education, Food habit, Hours_of_study. **Behavioral**: Communication skills, Social Network interactions, Medium of Education, Punctuality in submission of assignment, Sincerity, Mental ability skills, Decision making capability.

Moreover, the prediction functions (Output parameters) reported different types of outcomes like eligibility to admissions, transferability, retention / detention, drop-out status and grades in courses of study.

D. Degree of Attribute Relevance

Bayesian classification methods could also reveal the degree of attribute relevance as published in one of the works by Bharadwaj and Pal, while DT classification models used by Yadav and Pal highlighted some kind of graded attribute relevance infering 'GS' (Graduation _Stream) as the highest contributable attribute [11] [3].

Till date, all the student's performance evaluation tools do not declare the kind of effort they are still needed to be put, in order to pass through their ongoing courses of graduation study. None of the prediction models is able to optimize the depth, up to what extent these academic efforts must be enforced upon both at student level and at teacher level so that they can be predicted into the category of pass students in their forthcoming end-semester examinations.

III. Modeling Methodology

The proposed model was carefully developed with two-phase functionality. The first phase helped in arriving at predicted values of class variable i.e. 'at-risk' and 'above-risk' values of the test data instance (referring to students belonging to on-going course). This was achieved with posterior probability computations of Naïve Bayesian Classifier. The novelty of this approach is that the computations upon class labels were further used to arrive at relative relevance of attributes contributable to success / failure grades. The second phase extends the experiment by computing the degrees of involvement due to each of the above attribute in affecting the predicted risk-category (already computed above by NB posterior probabilities) of the students by generating precedence relations for those attributes. The authors do not give up here! The performance of the proposed feature-extraction-cum-ranking model is evaluated by extending the experiments with RELIEF method as benchmark [12].

IV. Experimental Evaluation

As the vast spectrum of attributes relevant to student databases can be collected easily, the authors have chosen the live student's data sets pursuing graduation course at one of their places of work for performing experiment of feature extraction. However, it was gradually realized that some of the attributes hardly affect the academic appraisal of the students due to their inherent and static nature.

In the proposed work, a method was devised to predict academic performance of the students in their on-going courses using training data patterns of passed-out batches purely by analyzing the attributes that contribute to their academic effort, once they are admitted to their current course of graduation study.

A. Parameter Settings and Data Pre-processing

The current experimental setup takes into account, students' attendance, internal assessment scores, assignment credit, and subject count (number of subjects, in which the student appeared in internal examination). These attributes, in turn act as external factors, which if enhanced are sure to improve end-semester results of on-going batch students.

The training dataset used to train the mining model included 87 students from three passed out batches of BCA course (Bachelor of Computer Applications taken up for experiments), as reported by College-Control-Unit of the institution. This was considered as input collection along with 20 test instances from the on-going batch of second year BCA students. The permissible domains upon the four attributes viz. students' attendance, internal assessment scores, assignment credit, and subject count were defined as: -7 to +7, 0..3, 0..10 and 0..10 respectively.

B. Attribute Precedence Relations

The computations begin with the classification of the test-tuples followed by attribute-wise fitness evaluation steps upon those tuples. The training data sets of 88 tuples from three passed-out batches of Second year BCA course were processed to compute prior probabilities of 'at risk' students. With the criterion set by CCU of the institution that, 'at-risk' students are prone to obtain less than 40% of aggregate score, the prior fit and unfit probabilities from the training data sets were found to be 0.76 and 0.24 for 'at-risk' and 'above-risk' students respectively. As the nature of the problem involves the appearance of four independent experimental parameters (x_1 to x_4), it was always appropriate to compute Naïve Bayesian posterior probabilities to compute the class labels,

International Journal of Advances in Computer Science & Its Applications
Volume 6 : Issue 1 [ISSN 2250-3765]
Publication Date : 18 April, 2016

namely 'at-risk' and 'above-risk' values. It can be recalled that the higher of these posterior probabilities computed for each test tuple t_i, pertaining to the current 2nd year batch: P(fit | $\{x_1, x_2, x_3, x_4\}$) and P(unfit | $\{x_1, x_2, x_3, x_4\}$) helped in deciding the predicted risk category of that each test-instance (t_i).

$$P(fit|\{x_1,x_2,x_3,x_4\}) = \frac{\sum_{i=1}^{4} p(fit).p\left(\frac{x_i}{fit}\right)}{\sum_{i=1}^{4} p(fit).p(\left(\frac{x_i}{fit}\right)+\sum_{i=1}^{4} p(unfit).p\left(\frac{x_i}{unfit}\right)} \tag{1}$$

$$average\_unfit(x_i,t_j) = \frac{\sum_{i=1}^{4} p(unfit).p\left(\frac{x_i}{unfit}\right)}{\sum_{i=1}^{4} p(fit).p\left(\frac{x_i}{fit}\right)+\sum_{i=1}^{4} p(unfit).p\left(\frac{x_i}{unfit}\right)} \tag{2}$$

The subsequent computations were extracted from the individual portions of the numerator components contributing to total conditional probabilities of fitness and unfitness. These numerator components of the (1) and (2) were revisited individually for computing average fitness (average_fit(x_i)) and average unfitness (average_unfit(xi)) of the students owing to each attribute as shown in (3) and (4).

$$average\_fit(x_i,t_j) = \frac{p(fit).p\left(\frac{xi}{fit}\right)}{\sum_{i=1}^{4} p(fit).p\left(\frac{x_i}{fit}\right)+\sum_{i=1}^{4} p(unfit).p\left(\frac{x_i}{unfit}\right)} \tag{3}$$

$$average\_unfit(x_i,t_j) = \frac{p(unfit).p\left(\frac{xi}{unfit}\right)}{\sum_{i=1}^{4} p(fit).p\left(\frac{x_i}{fit}\right)+\sum_{i=1}^{4} p(unfit).p\left(\frac{x_i}{unfit}\right)} \tag{4}$$

The relative attribute fitness evaluation step is performed with the underlying feature of NB classifier that the individual conditional probabilities upon each of the four attributes, x_1, x_2, x_3 and x_4 together contribute at classifying the 'above-risk' / 'at-risk' classification task. So their relative comparisons measured in terms of relative fitness / unfitness precedence for each test instance (students studying an on-going course) unfold the order in which the attributes contribute to his / her final performance status of being 'at-risk' or 'above-risk' at the end of that course tenure.

The above mentioned individual fitness probabilities namely average (fit(x_1)), average (fit(x_2)), average (fit(x_3)) and average (fit(x_4)) were computed owing to the four attributes from (3), which in turn could be laid in increasing order for comparisons; some of the test instances are illustrated in *table 1*. Similarly, the precedence relations could be compiled for average unfit (x_i) probabilities for the above mentioned four attributes from (4). These attribute precedence relations arranged in order of increasing fitness / unfitness, obtained from the above proposed FE modeling open the dimension of decision making tasks in the direction of individual student's counseling.

TABLE I. ATTRIBUTE PRECEDENCE RELATIONS OF FITNESS AND UNFITNESS (PROPOSED MODEL)

Tuple-ID	X1	X2	X3	X4	P(fit | x)	P(unfit | x)	Avg(x1 | fit)	Avg(x2 | fit)	Avg(x3 | fit)	Avg(x4 | fit)	Attribute Precedence Relations of Fitness	Attribute Precedence Relations of Unfitness
1	-7	4	1	5	0.21	0.79	0.10	0.00	0.08	0.00	x2<x4<x3<x1	x4 <x1 <x3<x2
6	-7	4	2	3	0.23	0.77	0.10	0.00	0.08	0.09	x2<x3<x4<x1	x1<x4<x3<x2
13	1	6	2	9	0.87	0.13	0.46	0.40	0.47	0.76	x2<x1<x3<x4	x4<x3<x1<x2
16	-1	5	1	3	0.36	0.64	0.00	0.17	0.08	0.09	x1<x3<x4<x2	x1<x2<x4 <x3
19	-7	5	1	3	0.34	0.66	0.10	0.17	0.08	0.09	x3<x4<x1<x2	x2 <x1<x4 <x3

C. Experiments with RELIEF

The performance evaluation of the proposed setup was decided to be performed using one of the popular feature extraction model 'RELIEF'. The popularity of the model is due to its increased accuracy, reduced time complexity, usage of simple statistical approach and enormous success achieved in practical applications.

D. RELIEF Parameters: 'Near-Hit' and 'Near Miss'

Here, the problem objective was solved with 'RELIEF' method, as it readily fits into two-class classification problem of predicting 'at-risk' level of students and identifying the precedence levels of attributes contributing to their academic performance.

Assuming the training instances are denoted by 'p' dimensional feature vector 'X' where p=4 for the current problem domain. The RELIEF algorithm makes use of p-dimensional Euclidean distance to select 'near-hit' and 'near-miss' instances from the training data set. For the current domain, 'Near-hit' (Z+) and 'Near-miss' (Z-) instances are defined as training instances (students from passed out batches) closest to the test instance but falling in 'pass' and 'fail' categories respectively. Revisiting the logic formulation by Kira and Rendall, the weights upon each of the participating attributes in the experimental feature vector were computed in (5) in order to lay attributes in increasing order of significance [16]. The authors also appreciate nearest-neighbor approach to find out 'nearest-hit' and 'nearest-miss' training instances to compute the weight updates as defined above.

$$\mathbf{w_i' = w_i - diff(x_i, near\text{-}hit_i)^2 + diff(x_i, near\text{-}miss_i)^2} \tag{5}$$

International Journal of Advances in Computer Science & Its Applications
Volume 6 : Issue 1 [ISSN 2250-3765]
Publication Date : 18 April, 2016

E. *Attribute Weight Computations*

For the above experimental setup to succeed, a minute but crucial issue needs attention to carefully devise the logistics for initializing weights over that p-dimensional feature vector attributes (referring to w_i component in (5)). The author conceptualizes two approaches for performing weight initialization step upon all the four experimental attributes for all twenty test instances for conducting performance evaluation experiments, some of which are tabulated in *table 2*.

1) Using prior probabilities of training instances: This p(.) value reveals the pattern from training data set, with what frequency (degree of chance); training instances bear the same value as that of the test instance.
2) Computing weights in normalized scale: The weights are mapped to normalized scale between 0 and 1 for all 'p' number of attributes in feature vector. For instance, $D_{attendance}$ = {-7, -6,….., -1,0,1,,…..6,7}; then for a test instance t_i, w(i, attendance) = (<attendance scale> + 8 /15).

TABLE II. WEIGHT INITIALIZATION COMPUTATIONS FOR PERFROMANCE EVALUATIONS WITH RELIEF

Parameters				Weights as prior probabilities				Normalized Weights			
x_1	x_2	x_3	x_4	wx_1	wx_2	wx_3	wx_4	wx_1	wx_2	wx_3	wx_4
-7	4	1	5	0.2	0.05	0.3	0.2	0.1	0.4	0.3	0.5
-7	4	2	3	0.2	0.05	0.4	0.03	0.1	0.4	0.7	0.3
1	6	2	9	0.04	0.3	0.43	0.03	0.6	0.6	0.7	0.9
-1	5	1	3	0.01	0.1	0.3	0.03	0.5	0.5	0.3	0.3
-3	5	1	8	0.02	0.1	0.3	0.2	0.3	0.5	0.3	0.8
-7	5	1	3	0.2	0.1	0.3	0.03	0.1	0.5	0.3	0.3

V. Results and discussion

The experiments was further extended in the direction of obtaining attribute precedence relations using RELIEF approach by computing weight updates due to (5) upon both the methods of weight initialization. The attribute precedence relations generated by 'RELIEF' heuristics acted as benchmark to find the accuracy of attribute precedence relations generated by 'proposed FE approach' and were interpreted for discussing the counseling directions and priorities of each of the student put up as test instance.

A. *Attribute Precedence Comparisons (Prior probabilities as weights)*

A rationalized thought was hypothesized to find the similarity between the corresponding sets of attribute precedence relations. The mining objective may be recalled as identifying up to what level, the attributes affect the academic performance levels with the baseline fact that all the attributes bear partial or strong relevance to their academics. Change in attribute precedence by one position only hardly changes the degrees of significance of the attributes and hence, the direction of counseling that student instance. The above opinion helped the work group in making use of both, total and partial precedence match patterns of attribute precedence (having precedence deviations by 1 position) as evaluation parameters. *Table 3, columns 7 and 8* compare the precedence relations showing computed graded attribute relevance due to both RELIEF (Method I) of weight initialization as well as proposed FE logistics as described in previous section. As a result, the performance of the model showed encouraging result with a sweep of model accuracy (82%) due to this method of RELIEF weight initialization (Method I)[a] .

B. *Attribute Precedence Comparisons (Normalized weights)*

Considering the second method of RELIEF weight formulations mapped in normalized scale, the attribute precedence relations so generated, were again compared with those obtained from proposed FE model (*table 4, columns 7 and 8*) using similar evaluation steps. This time, when the total and partial attribute precedence match counts were computed on four point scale, the performance of the model showed even better results (accuracy 83%) due to this RELIEF weight initialization step [b].

TABLE III. ATTRIBUTE PRECEDENCE COMPARISONS WITH INITIALIZED RELIEF WEIGHTS (METHOD A)[A]

Tuple-ID	Pass/Fail	Wx1'	Wx2'	Wx3'	Wx4'	RELIEF method	Proposed Method
1	1	0.2	0.1	0.3	0.2	x2<x4=x1<x3	x2<x4<x1<x3
2	1	0.1	-0.9	0.4	-1.0	x4<x2<x1<x3	x2<x1<x4<x3
3	1	12.0	1.1	0.3	-9.0	x4<x3<x2<x1	x4<x2<x3<x1
4	1	0.0	9.3	3.2	9.2	x1<x3<x4<x2	x4<x2<x1<x3
5	1	4.4	0.8	1.3	0.1	x2<x4<x3<x1	x4<x3<x2<x1
6	0	0.2	-0.9	0.4	-1.0	x4<x2<x1<x3	x4<x2<x1<x3
7	1	4.0	0.3	-0.7	3.2	x3<x2<x4<x1	x3<x2<x1<x4
8	0	0.1	-0.9	0.4	-1.0	x4<x2<x1<x3	x2<x1<x4<x3
9	1	-1.7	1.0	0.2	-1.7	x1=x4<x2<x3	x4<x2 <x1<x3
10	1	0.4	1.2	1.2	1.0	x1<x4<x2=x3	x2<x1<x4<x3
11	1	0.4	0.3	0.4	0.0	x4<x2<x1=x3	x2<x1<x3<x4
12	1	3.6	2.5	0.3	3.8	x3<x2<x1<x4	x3<x2<x1<x4
13	1	3.0	2.0	0.2	0.0	x4<x2=x3<x1	x2<x3<x4<x1
14	1	4.0	-3.9	1.4	3.2	x2<x3<x4<x1	x2<x3<x1<x4
15	1	-3.5	-0.5	0.3	-0.5	x2<x4<x1<x3	x2<x3<x1<x4
16	1	12.0	1.1	0.3	-9.0	x4<x3<x2<x1	x4<x2<x3<x1
17	1	16.0	3.3	0.3	4.2	x2=x3<x4<x1	x4<x2<x3<x1
18	0	0.2	0.1	0.3	0.2	x3<x2<x1<x4	x3<x2<x4<x1
19	1	0.1	-0.9	0.4	-1.0	x4<x1=x2<x3	x4<x2<x1<x3
20	0	12.0	1.1	0.3	-9.0	x2<x1<x4<x3	x1<x4<x2<x3

a. Weighted Accuracy of attribute precedence=1 * 9/20 + .75 * 9/20 + .5 * 01/20 + .25 * 0/20= 82%

TABLE IV. ATTRIBUTE PRECEDENCE COMPARISONS WITH NORMALIZED RELIEF WEIGHTS (METHOD II)[B]

Tuple-ID	Pass/Fail	Wx1'	Wx2'	Wx3'	Wx4'	RELIEF method	Proposed Method
1	1	0.1	0.4	0.3	0.5	x1<x3<x2><x4	x2<x4<x1<x3
2	1	0.2	-.3	0.7	-0.4	x4<x2<x1<x3	x2<x1<x4<x3
3	1	12.5	0.5	0.3	-4.7	x4<x3<x2<x1	x4<x2<x3<x1
4	1	0.3	0.7	-.0	8.8	x3<x1<x2<x4	x4<x2<x1<x3
5	1	5.0	0.7	0.3	-0.5	x4<x3<x1<x2	x4<x3<-x2<x1
6	0	0.1	-.6	0.7	-0.7	x4<x2<x1<x3	x4<x2<x1<x3
7	1	4.6	-.4	0.3	0.7	x2<x3<x4<x1	x3<x2<x1<x4
8	0	0.2	-.5	0.7	-0.6	x4<x2<x1<x3	x2<x1<x4<x3
9	1	-1.6	-0.5	1.0	-2.7	x4<x1<x2<x3	x4 <x2<x1<x3
10	1	1.0	0.8	2.0	1.9	x2<x4<x1<x3	x2<x1<x3<x4
11	1	1.0	0.7	0.7	0.9	x2=x3>x4<x1	x2<x1<x3<x4
12	1	4.2	0.6	-0.3	4.8	x3<x2<x1<x4	x3 <x2<x1<x4
13	1	3.6	0.6	-0.3	0.9	x3<x2<x4<x1	x2<x3<x4<x1
14	1	4.6	-3.5	1.7	3.7	x2<x3<x4<x1	x2<x3<x1<x4
15	1	3.4	-2.6	0.3	-2.7	x1<x2<x4<x3	x2<x3<-x1<x4
16	1	12.5	0.5	-0.7	-8.7	x4<x2<x3<x1	x4<x2<x3<x1
17	1	16.5	-0.2	0.6	-3.0	x4<x2<x3<x1	x4<x2<x3<x1
18	0	4.3	1.5	0.3	5.8	x1<x3<x2><x4	x2<x4<x1<x3
19	1	0.1	0.8	0.3	-0.7	x4<x2<x1<x3	x2<x1<x4<x3
20	0	-3.3	-3.5	1.0	-0.5	x4<x3<x2<x1	x4<x2<x3<x1

b. Weighted Accuracy of attribute precedence=1 * 11/20 + .75 * 6/20 + .5 * 2/20 + .25 * 0/20= 83%

VI. Conclusion and Future work

The favorable accuracies obtained due to comparisons of proposed FE model with both RELIEF-A and RELIEF-B models ensured its consistent performance for generating attribute precedence levels upon the mentioned 4-attribute schema. Consequently, the attributes positioned with greater fitness levels were found more contributing to pursue effective student's counseling in order to improve their academic appraisals. The proposed framework came out with many valuable decision-making tips that helped the management to take needful pre-emptive actions in an attempt to upgrade their grades in forthcoming examinations. The experiments shall be extended by increasing the feature-vector length (p=6 or 9) to study the performance of the model there after.

Acknowledgment

The authors acknowledge Department of Computer Science, Sai Mahavidyalaya, Sector-6, Bhilai, Chhattisgarh, India for providing live data sets of pre-final year students pursuing a technical course (Bachelors of Computer

Applications) affiliated to Pt. Ravi Shankar Shukla University, Chhattisgarh, India.

References

[1] U. K. Pandey and S. Pal, "Data Mining: A prediction of performer or underperformer using classification", Int. J. Computer Science and Information Technologies, vol. 2, pp. 686-689, 2011.

[2] U. K. Pandey and S. Pal S, "Data Mining View on Classroom Teaching Language",Int. J. Computer Science Issues, vol. 8, No. 2, , pp. 277-282, 2011.

[3] Yadav S.K. and Pal S. (March 2012), "Data Mining Application in Enrollment Management: A Case Study", Int. J. Computer Applications, vol. 41, No. 5, pp. 1-6, March 2012.

[4] M. N. Quadril and N. V. Kalyankar, "Drop Out Feature of Student Data for Academic Performance Using Decision Tree Techniques", Global J. Computer Science and Technology, vol. 101, No. 2, pp. 2-5, 2010.

[5] S. K. Yadav, B. Bharadwaj and S. Pal, "Mining Education Data to Predict Students' Retention: A Comparative Study", vol. 10, No. 2, , pp. 113-117, 2012.

[6] B. M. Bidgoli and W. F. Punch, "Using Genetic Algorithms for Data Mining Optimization in an Educational Web-Based System", Genetic Algorithms Research and Application Group, LNCS Notes 2724, pp. 2252-2263, 2003.

[7] K. Shymala and S. P. Rajgopalan, "Data Mining Model for a Better Higher Educational System", Information Technology J., vol. 5, No. 3, pp. 560-564, 2006.

[8] M. Fire, G. Katz, Y. Elovici, B. Shapira and L. Rokach, Telecom Innivation Laboratories and Information Systems Engineering Department, 2012.

[9] S. B. Kotsiantis, C. Pierrakeas, I. D. Zaharakis and P. E. Pintclas, "Preventing student dropouts in distance learning using machine learning techniques", pp. 267-274, ISBN 3-540-40803-7, 2003[Proc. 7th Int. Conf. Knowledge-Based Intelligent Information and Engineering Systems].

[10] E. Osmanbegovic and M. Suljic "Data Mining Approach for Predicting Student Performance", vol. 10, No. 12, Economic Review J. Economics, pp. 3-12, May 2012.

[11] B. K. Bharadwaj and S. Pal, "Data Mining: A Prediction for Performance Improvement using Classification", Int. J. Computer Science and Information Security, vol. 9, No. 4, pp. 136-140, 2011.

[12] K. Kira and L. A. Rendell L.A., "The Feature Selection Problem: Traditional Methods and a new Algorithm", Proc. AAAI, pp129-133 , 1992.

About Author (s):

Mamta Singh (MCA, Maharshi Dayanand University, 2005, M.Phil, Periyar University) is a Research Scholar working on "Prediction Of Academic Performance Of Students Using Machine Learning Techniques".

Dr. (Prof.) Jyoti Singh (MCA, Banasthali Vidyapith, Rajasthan, 1990, Ph.D., Pt. Ravi Shankar Shukla University, Raipur, 2007) is a Life member of "The Indian Society for Technical Education" and approved PhD Supervisor in CSVTU Bhilai, Dr. C V Raman University, Chhattisgarh supervising eight research scholars as on date.

Dr. (Prof.) Arpana Rawal (M.Tech. Computer Technology, Pt. Ravi Shankar Shukla University, Raipur, 2003, Ph.D. Chhattisgarh Swami Vivekanand University, 2013 is a Life member of "The Indian Society for Technical Education" as well as Computer Society of India". Her areas of research interest are Intelligent Systems, Data Mining, Machine Learning using text mining and Natural Language Processing.