Naïve Bayesian driven Feature Extraction Model: A Comparative Study

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

Abstract—The ongoing research on innovative Feature Extraction modeling using Naïve Bayesian computations have catered to meaningful research directions in qualitative teaching learning environments. The n-attribute Feature Extraction Modeling experiments initially involve the appearance of four independent academic parameters that are totally changeable and explicitly specified, while the students pursue their academic semesters. The research experiment was also extended by adding more parameters to formulate Six-Attribute and Nine-Attribute Feature Extraction Models. The model accuracies obtained from the three variants of feature extraction models were compared. All the mentioned FE models also provided sound interpretations on the relative academic effort offset yet toput by the students, so as to come out of 'risk' category during their forthcoming, endsemester/annual examination.

Keywords — Feature Vector, Feature Extraction model, Machine learning, Naïve Bayesian learning, Optimal Attribute Precedence Relations, RELIEF Feature - extraction model, Model Accuracy.

1 INTRODUCTION

Educational Data Mining is a constantly growing research area since early 1990s that is being gradually adopted by educational centers in order to implement more effective pedagogical strategies and help instructors to validate and evaluate teaching-learning processes so that suitable feedbacks-cum-recommendations can be obtained by students about their status of attaining those learning objectives in respective courses. In EDM research timeline decision making dimensions are explored in three perspectives that are Students, Educators and Managerial strategies and help instructors to validate and evaluate teaching-learning processes so that suitable feedbacks-cum-recommendations can be obtained by students about their status of attaining those learning objectives in respective courses.

Some of the data mining tasks that can be performed in the interests as viewed from various students' perspective are seeking recommendations on improving their learning levels, suggesting topic-learning path pruning, getting alerts on retention levels in the on-going course, online or individual counseling in case the student is predicted 'at-risk'. The purpose of current research to develop a Naïve Bayesian prediction cum ranking model that is acts as quantitative recommender for

Potentially weak students who are likely to fail, so that educators can suggest appropriate remedial actions in form of relative academic efforts. Some of already explored EDM tasks till date include: analysis and visualization of data, providing feedback for support teaching domain, recommendation for students, prediction modeling on student performance, student cat-egory and behavior modeling, social network analysis in online learning enviornments, building topic and concept maps and lastly content based-planning and scheduling courseware material [1][2][3][4][5].

2 FEATURE EXTRACTION MODELING IN EDM

No doubt, the prediction model highly depends on the choice of selection of input parameters used in the mining task. The same rule of thumb applies to student at-risk prediction modeling that plays a crucial role in developing and improving students' academic appraisal and hence forth high quality recommender systems well in time. Feature selection or subset selection is a preprocessing step commonly used in machine learning. The aim of feature selection is to minimize the large set of features. In Feature Extraction approach, all irrelevant and redundant features are removed and a subset of the features available from the data-sets is selected as input parameters to the learning algorithm. The elimination of irrelevant and redundant information is said to improve the quality of learning and also accurcy of learning model. Till date a multitude of feature selection models have been developed that aim to extract those combinations of attributes yielding improved classification accuracy.

2.1 Feature Vector Ranking

Kira and Rendell (1992) designed RELIEF algorithm that assigned a relevant weight to each attribute of feature vector by computing difference between the selected test instance and as nearest hit and nearest miss training instances. John et al. have attempted to extract features sub-sets using supervised machine learning with induction methods (ID3 and C4.5). For the first time, hypothesized graded notions of relevance (strongly relevant, weakly relevant, were introduced. Their study investigated the possibility of improving prediction accuracy or decreasing the size of the structure without losing prediction accuracy [6] [7].

Sun and Wu (2008), during their in-depth study on feature selection methods proved that RELIEF is the most successful algorithm that solves a convex optimization problem with a margin based objective function. As, it was observed that the RELIEF model couldn't filter out redundant attributes as well as weakly relevant ones, this motivated the author to provide variant logistics to the approach,[8].

Sewell. M (2007) did survey on feature extraction and presented a concept that a subset contains the least number of dimensions contributes to greater accuracy. They categorized feature selection methods into three types as complete, heuristic and random. Their study provided built-in mechanism for automatically removing irrelevant features during learning [9].

A survey reporting by Aziz A. A. et al (2013) that focused on three elements needed to make prediction on student' academic performance: parameters, methods and tools. Naïve Bayes classifier was used to extract pattern using the data mining WEKA tool. Their study was preceded with the implementation of the framework for predicting SAP (student academic performance). After normalizing department database, few set of feature parameter were selected. The framework model consists of data collection, data transformation, pattern extraction, prototype development phases [10].

Observing the methodology trends used for building learning model the current research work on feature extraction task is also based onthe current research trends that aims to to develop model with the objective of the feature of the proposed learning model, (Naïve Bayes driven learning model), as it works well both for prediction as well as utilizing the prediction results in gearing up the student's potential during their on-going period of studies

3 Data Collection and Modelling Parameter

The presented EDM framework begins with setting of input Parameters that reflect the students' performance directly or indirectly from the instant, the students take up their courses. Initially four attributes viz. student attendance, assignment Credit, internal score and subject count have been taken to begin with the first experiment and so has been given the name, "Four-Attribute FE-Model". In this experiment, all taken attributes were dynamic or changeable in nature. By the term dynamicity we mean that students can improve upon above mentioned attributes their by putting more efforts.

Current experimental approach is extension of previous prediction model (Four-FE Model). In subsequent experiments, attribute schema increased with two more attributes by taking into account, Laboratory Credit and Previous year Percentage of students and model has been given a name Six-Attribute Model.

Data collection of values for laboratory credit and previous year percentage (i.e. percentage in first year of graduation course) is credited to the data source from University Tabulation Register (TR sheets). 'Laboratory Credit' is changeable in nature that means, after gaining an insight on course attribute amidst the academic session, the students have the scope to increase the value of their Laboratory Credit by giving extra efforts during their remedies of practical sessions. Past Year Percentage is static in nature because even after arriving at the attribute precedence relations on such attribute, the value pertains to last year's declared result and hence non-changeable, however, is presumed to play a major role in the experiment for predicting results of the current academic session.

Yet another set of add-ons were made by taking into account, three more attributes over above defined Six Attribute model i.e. Medium of Study (i.e. English or Hindi), Percentage of Higher Secondary and Living Location (Rural, Semi urban, Urban) values were collected from the students, Admission and enrollment form, students when they had taken admission in their courses. All these three additional attributes were found indirectly contributable to research objective stated in the setup. Two of the attributes, language as well as percentage obtained in higher secondary course is non changeable in nature but were included in the experiment, owing to certain past referred studies that had considered these parameters in constructing knowledge model to evaluate students academic performance in on-going courses thereafter. Living location is also subsumed to be indirectly contributable as it cannot be violated that the commuting distance and time as well as demographics do affect the students hours of study and hence their academic performance.

4 Feature extraction-cum-Ranking Model / Modelling approach

The whole experimental setup is outlined in three successive stages: Four-Attribute Feature-Extraction (FE) model, Six-Attribute Feature-Extraction (FE) model and Nine-Attribute Feature-Extraction (FE) model. A paper has already been published describing 'four-attribute FE model' and comparing theaccuracy results with an equivalently built FE model based on RELIEF feature-extraction method [10] [11].

In each set of experiments the computation begins with the classification of the test-tuples into fit and unfit predictions using Naïve Bayesian Classifier. This was followed by attribute-wise fitness evaluations upon those tuples. The training data sets of 87 tuples from three passed-out batches of second year bachelor of Computer Application course were processed to compute prior probabilities of 'at risk' students The data of eighty seven students of college is collected who had appeared for the second year of graduation course in the past three consecutive academic sessions and stored as training data sets. Similarly twenty instances of test data were collected from the ongoing academic session of the same course.2013-14.

As the nature of the problem initially involves the appearance of four independent experimental parameters (x_1 to x_4), it was always appropriate to compute Naïve Bayesian posterior probabilities over class labels, namely 'at-risk' and 'above-risk' values for Four Attribute FE model. It can be recalled that the higher of these posterior probabilities computed for each test tuple ti, pertaining to the current second year batch: P (fit | $\{x_1, x_2, x_3, x_4\}$) and P (unfit | $\{x_1, x_2, x_3, x_4\}$) helps in deciding the predicted risk category of that each test-instance (ti). For instance, p (fit | [.]) is more than p (unfit | [.]) for a student, then he / she belongs to"above risk" category, as shown in figure 4.1.

ISSN 2229-5518

tuple_id	p(fit|ti)	p(unfit|ti)	pass/fail
1	0.82	0.18	1
2	0.90	0.10	1
3	0.82	0.18	1
4	0.99	0.01	1
5	0.93	0.07	1
6	0.74	0.26	1
7	0.96	0.04	1
8	0.90	0.10	1
9	0.82	0.18	1
10	0.99	0.01	1
11	0.98	0.02	1
12	0.97	0.03	1
13	0.97	0.03	1
14	0.95	0.05	1
15	0.74	0.26	1
16	0.82	0.18	1
17	0.95	0.05	1
18	0.93	0.07	1
19	0.76	0.24	1
20	0.82	0.18	1

Figure 4.1 Predicted At_Risk / Above_Risk value in Four-Attribute Feature Extraction Model

Six-Attribute Feature Extraction (FE) model used six attributes of analysis while, Nine-FE model used nine experimental attributes as shown in TABLE 1. The posterior probabilities of fitness unfitness for both the extended FE models are defined in the expressions 4.1 to 4.4.

TABLE 4.1 EXPERIMENTAL ATTRIBUTES FORMULATION

Sl. No.	Name of the Parameter	Parameter Description	Domain Values	Domain Value
X_1	Attendance	Student's attendance from July to January. Minimum 70% attendance is compulsory.	0..100%	+7,+5 ,+3, +1, 0, -7,-5, -3, -1}
X_2	Assignment Credit	Assignment Credit were given by teachers in theory subject	0..10 (Marks)	0..10
X_3	Internal Score	Performance in three internal unit tests	0….100%	0..3
X_4	Subject count	Student appear in how many internal paper out of 10	0…...100%	0..10
X_5	Lab Credit	Laboratory Score in three practical subjects	40%……100%	1….10
X_6	Past Percentage	Previous Exam Grade or score	40%….100 %	1….10
X_7	Higher Secondary Percent	Percentage Scores in class12th Board Examination	33%….100 %	2…4
X_8	Medium of Study	Medium of instructions in class room	English/Hindi	1-2
X_9	Living Location	Residential status of Students	Rural-1, Urban-2, Sub-urban-3.	1-3

$$P(fit \mid \{x_1,x_2,x_3,x_4,x_5,x_6\}) = \frac{\sum_{i=1}^{4} p(fit).p\left(\frac{x_i}{fit}\right)}{\sum_{i=1}^{4} p(fit).p(\left(\frac{x_i}{fit}\right)+\sum_{i=1}^{4} p(unfit).p\left(\frac{x_i}{unfit}\right)} \;\ldots\ldots\ldots \quad (4.1)$$

$$P(unfit \mid \{x_1,x_2,x_3,x_4,x_5,x_6\}) = \frac{\sum_{i=1}^{4} p(unfit).p\left(\frac{x_i}{unfit}\right)}{\sum_{i=1}^{4} p(fit).p\left(\frac{x_i}{fit}\right)+\sum_{i=1}^{4} p(unfit).p\left(\frac{x_i}{unfit}\right)} \;\ldots\ldots\ldots \quad (4.2)$$

$$p(fit \mid \{x_1,x_2,x_3,x_4,x_5,x_6,x_7,x_8,x_9\}) = \frac{\sum_{i=1}^{4} p(fit).p\left(\frac{x_i}{fit}\right)}{\sum_{i=1}^{4} p(fit).p(\left(\frac{x_i}{fit}\right)+\sum_{i=1}^{4} p(unfit).p\left(\frac{x_i}{unfit}\right)} \ldots \quad (4.3)$$

$$P(unfit \mid \{x_1,x_2,x_3,x_4,x_5,x_6,x_7,x_8,x_9\}) = \frac{\sum_{i=1}^{4} p(unfit).p\left(\frac{x_i}{unfit}\right)}{\sum_{i=1}^{4} p(fit).p\left(\frac{x_i}{fit}\right)+\sum_{i=1}^{4} p(unfit).p\left(\frac{x_i}{unfit}\right)} \;. \quad (4.4)$$

The subsequent computations form the feature-vector ranking step in each of the feature extraction modes. Here, the individual portions of the summative numerator components of expressions 4.1 to 4.4 were extracted. Each component pertained to single attribute, reflecting the posterior effect of that attribute on fitness parameter. A novel thought of evaluating relative attribute fitness / unfitness was implemented with the underlying feature of NB classifier that the individual conditional probabilities upon each of the 'n' attributes x_1 ttogether contribute at classifying the 'above-risk' / 'at-risk' classification task. The comparisons among such individual portions helped in ranking the attributes of experimental feature vector.

In order to generate the precedence order of these experimental attributes, the individual numerator components of the 4.1 to 4.4 were revisited and used for computing average fitness (average_fit(x_i,t_j)) and average unfitness (average_unfit(xi, tj)) of the students owing to each attribute as shown in expression (4.5) and (4.6) respectively.

$$average\_fit(x_i, t_j) = \frac{p(fit).p(\frac{xi}{fit})}{\sum_{i=1}^{4} p.(fit)p\left(\frac{x_i}{fit}\right)+\sum_{i=1}^{4} p(unfit).p(\frac{xi}{unfit})} \ldots\ldots\ldots \quad (4.5)$$

$$average\_unfit(x_i, t_j) = \frac{p(unfit).p(\frac{xi}{unfit})}{\sum_{i=1}^{4} p(fit).p\left(\frac{x_i}{fit}\right)+\sum_{i=1}^{4} p(unfit).p(\frac{x_i}{unfit})} \ldots\ldots \quad (4.6)$$

ISSN 2229-5518

Name of Stud	Tuple_id	p/f	precedence fitness					
Adarsh Shrivast	1	1	0.17-x2	0.24-x4	0.27-x1	0.31-x3	0.32-x5	0.42-x6
Anshul Lanjawre	2	1	0.23-x5	0.25-x2	0.32-x1	0.32-x4	0.36-x6	0.36-x3
Chandan Kumar S	3	0	0.00-x4	0.16-x5	0.21-x6	0.25-x2	0.31-x3	0.42-x1
Deepesh Ganjir	4	1	0.16-x5	0.40-x4	0.40-x2	0.42-x1	0.42-x3	0.42-x6
Ku. Jigyasha	5	1	0.24-x4	0.28-x6	0.31-x3	0.32-x5	0.40-x2	0.41-x1
Manoj Kumar	6	0	0.00-x4	0.16-x5	0.17-x2	0.21-x6	0.27-x1	0.36-x3
Md. Majhar	7	1	0.31-x3	0.32-x5	0.33-x2	0.36-x6	0.42-x1	0.42-x4
Mona Mandavi	8	1	0.25-x2	0.32-x1	0.32-x4	0.32-x5	0.36-x3	0.42-x6
Nandita Sardar	9	1	0.00-x4	0.23-x5	0.25-x2	0.32-x1	0.42-x3	0.42-x6
Neha Khandelwal	10	1	0.32-x5	0.40-x2	0.41-x1	0.42-x3	0.42-x4	0.42-x6
Nidhi Khandelwa	11	1	0.28-x6	0.36-x3	0.40-x5	0.40-x2	0.41-x1	0.42-x4
Pratima Soni	12	1	0.28-x6	0.33-x2	0.36-x3	0.40-x4	0.40-x5	0.42-x1
Promod Kumar Sa	13	1	0.32-x5	0.33-x2	0.36-x3	0.42-x1	0.42-x4	0.42-x6
Ruchika Pandey	14	1	0.23-x5	0.25-x2	0.36-x3	0.42-x1	0.42-x4	0.42-x6
Saleya Khatoon	15	0	0.00-x4	0.17-x2	0.23-x5	0.31-x3	0.32-x1	0.42-x6
Sanju Patel	16	1	0.00-x4	0.25-x2	0.31-x3	0.40-x5	0.42-x1	0.42-x6
Santosh Kumar	17	1	0.16-x5	0.31-x3	0.33-x2	0.36-x6	0.40-x4	0.42-x1
Suraj Tripathi	18	1	0.16-x5	0.25-x2	0.31-x3	0.36-x6	0.40-x4	0.42-x1
T. anita Soni	19	0	0.00-x4	0.23-x5	0.25-x2	0.27-x1	0.31-x3	0.36-x6
Yogendra shyam	20	0	0.14-x1	0.21-x6	0.23-x5	0.24-x4	0.25-x2	0.36-x3

Figure 4.2 Attribute Precedence Relations of Fitness Six-Attribute FE Model

5 Feature Vector Ranking with RELIEF

Further experiments were carried out in the direction of obtaining accuracy of attribute precedence relations from the Bayesian driven Hybrid model for all the attributes in p-dimensional feature vector (p=6 and p=9). 'RELIEF' heuristics was used as benchmark to find the model's accuracy and was interpreted for discussing the counseling directions and academic effort-priorities of each of the student put up as test instance.

The comparisons of the attribute precedence relations between the Bayesian driven FE model and RELIEF FE model were done by computing similarity between the sets of attribute precedence relations obtained through the three mentioned FE models as well as comparing by RELIEF method.The weight update operation was performed upon each of the participating attributes in the experimental feature vector. These updated attribute weights act as rank values of the attributes when sorted in increasing order of relevance. The author also appreciate nearest-neighbor approach to find out 'nearest-hit' and 'nearest-miss' training instances to compute the weight updates as defined in expression 5.1. It is obvious that the 'near-hit' and 'near-miss' training instances shall differ for each test instance, as compared to those extracted for four-attribute FE model.

$$\mathbf{w_i' = w_i - diff(x_i, near\text{-}hit\text{-}instance_i)^2 + diff(x_i, near\text{-}miss\text{-}instance_i)^2} \cdots\cdots\cdots (5.1)$$

5.1 Performance Evaluations of Six-Attribute FE Model

The performance evaluation setup of six attribute FE model had (n = 87) training instances, each denoted by 'p' dimensional feature vector 'X' where p=6 for the current problem domain. The RELIEF algorithm makes use of four-dimensional euclidean distance to select 'near-hit' and 'near-miss' instances from the training data set. For the current domain,' Near-hit' and 'Near-miss' instances were defined as training instances (students from passed out batches) closest to the test instance but falling in 'PASS' and 'FAIL' categories, symbolized as Z+ and Z- respectively. If the test instance x_j is predicted as a 'PASS' instance (i.e. the student obtaining 'above-risk' status) then near-positive training neighbor with 'PASS' class label (x_j) is assigned as Z+ and nearest negative training neighbor (x_j) is assigned with Z- component. Instead, if the test instance falls into 'at-risk' predicted status, its nearest negative training neighbor holding 'FAIL' class label is assigned Z+ and nearest positive training neighbor with 'PASS' class label is assigned Z- component. The 'near-hit' (Z+) and 'near-miss'(Z-) training instances were used for the experimental test data set.

It may be noted that the two conceptualized weight initialization approaches used were weights in normalized scale and weights with prior probabilities, designated as methods I and II respectively.The details of applying such heurestics to weight to weight initialization is already described by author in one of their published work [12].The weight initialization step contributes to computation of first component in the weight update expression 5.1. In this way equivalent Attribute Precedence Relations (APRs) were found by computing the weight updates for all attributes in Six-Attribute FE model and Nine-Attribute FE models.

The Figures 5.1 and 5.2 show the comparisons of the attribute precedence relations between the Bayesian driven FE model and RELIEF FE model by computing similarity between the sets of attribute precedence relations obtained through the above mentioned models.

AN22

	A	C	D	E	F	G	H	J	K	L	M	N	O	P	U	AB	AC	AD	AE	AF	AG	AH	A
1	S No	1	2	3	4	5	6		1	2	3	4	5	6		1	2	3	4	5	6	Per %	%
2	1	x6	x5	x2	x1	x4	x3		x2	x4	x1	x3	x5	x6		0	0	0	1	0	0	16.67	%
3	2	x4	x2	x5	x3	x6	x1		x5	x2	x1	x4	x6	x3		0	1	0	0	1	0	33.33	%
4	3	x4	x5	x2	x3	x6	x1		x4	x5	x6	x2	x3	x1		1	1	1	1	0	1	83.33	%
5	4	x3	x5	x2	x6	x4	x1		x5	x4	x2	x1	x3	x6		0	1	1	0	0	0	33.33	%
6	5	x4	x5	x6	x2	x3	x1		x4	x6	x3	x5	x2	x1		1	0	1	1	0	1	66.67	%
7	6	x5	x4	x2	x6	x1	x3		x4	x5	x2	x6	x1	x3		1	1	1	1	1	1	100	%
8	7	x2	x4	x3	x5	x6	x1		x3	x5	x2	x6	x1	x4		0	0	0	0	1	1	33.33	%
9	8	x4	x2	x6	x5	x3	x1		x2	x1	x4	x5	x3	x6		0	1	0	1	1	0	50	%
10	9	x4	x2	x6	x3	x5	x1		x4	x5	x2	x1	x3	x6		1	1	0	1	0	0	50	%
11	10	x5	x2	x1	x6	x4	x3		x5	x2	x1	x3	x4	x6		1	1	1	0	1	0	66.67	%
12	11	x4	x6	x2	x3	x5	x1		x6	x3	x5	x2	x1	x4		0	1	1	0	0	1	50	%
13	12	x3	x6	x2	x5	x1	x4		x6	x2	x3	x4	x5	x1		0	1	1	1	1	0	66.67	%
14	13	x5	x3	x4	x2	x6	x1		x5	x2	x3	x1	x4	x6		1	1	0	0	1	0	50	%
15	14	x5	x2	x6	x3	x4	x1		x5	x2	x3	x1	x4	x6		1	1	0	1	1	0	66.67	%
16	15	x4	x2	x6	x5	x3	x1		x4	x2	x5	x3	x1	x6		1	1	0	1	1	1	83.33	%
17	16	x4	x3	x2	x5	x6	x1		x4	x2	x3	x5	x1	x6		1	1	1	1	1	1	100	%
18	17	x4	x2	x5	x6	x3	x1		x5	x3	x2	x6	x4	x1		0	1	0	1	0	1	50	%
19	18	x6	x5	x3	x2	x4	x1		x5	x2	x3	x6	x4	x1		0	1	1	0	1	1	66.67	%
20	19	x5	x4	x1	x3	x2	x6		x4	x5	x2	x1	x3	x6		1	1	1	1	0	1	83.33	%
21	20	x2	x5	x1	x4	x3	x6		x1	x6	x5	x4	x2	x3		0	1	0	1	1	0	50	%
22																						Total	
23								Total no of tupple of 25 %														1	
24								Total no of tupple of 50 %														9	
25								Total no of tupple of 75 %														5	
26								Total no of tupple of 100 %														5	
27								Total no of tupple of 0 %														0	
28								Average Accuracy														60	%

Figure 5.1 Attribute Precedence comparisons for 6-Attribute Bayesian Driven FE Model VsRELIEF (Normalized) FE Model, (Method I)

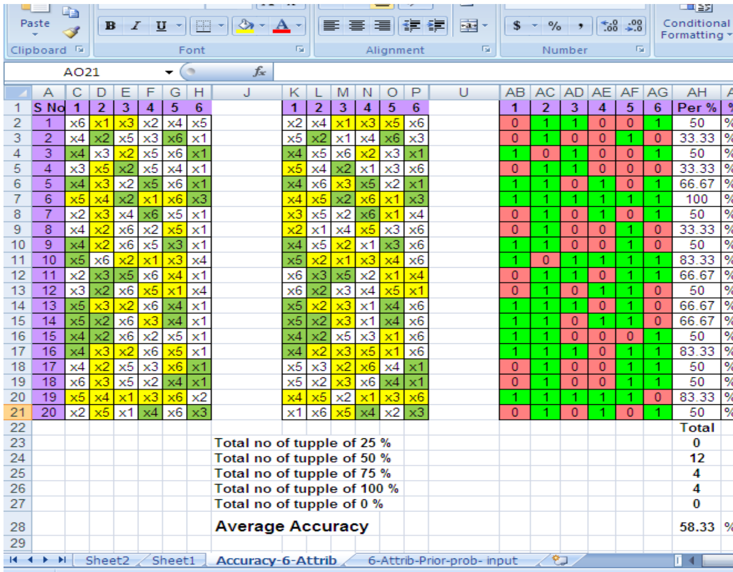

Fig 5.2 Attribute Precedence comparisons for 6-Attribute Bayesian Driven FE Model Vs RELIEF (Prior-probabilities) FE Model, (Method II)

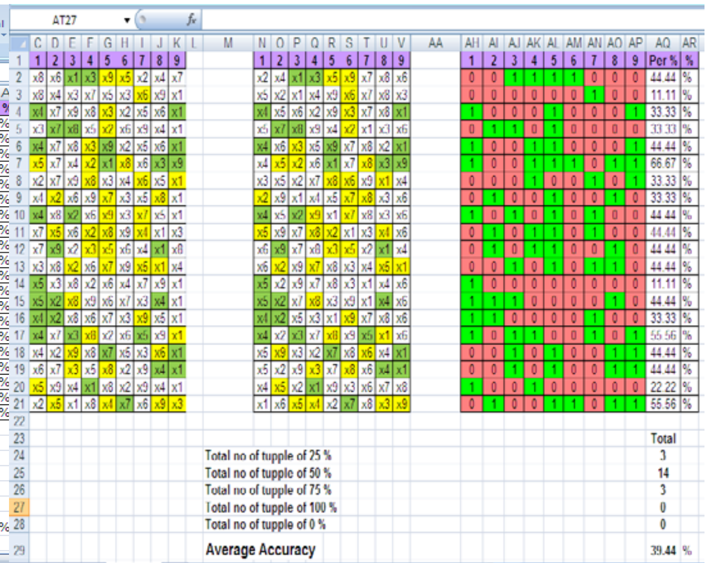

Fig 5.3 Attribute Precedence comparisons for Nine-Attribute Bayesian Driven FE Model Vs REIEF (Normalized) FE Model, (Method I)

As can be seen in figure 5.1 the dark highlighted attribute-relations are five in number that exhibit the total match in precedence while shallowly highlighted attributes exhibit partial precedence match instances: i. e. one relation method in ratio of 25%, nine relations were matched 50% and five relations were found 75% matched.

On the Controry, APRs obtained by RELIEF method II exhibited following the match % ratios for six-attribute precedence relations: ([NIL relation, 25%] [twelve relations, 50%] [four relations, 75%] [four relations, 100%]).

5.2 Performance Evaluations of Nine-Attribute FE Model

The third experimental setup proceeded by adding three additional attributes to the FE model described in previous section The first addition was higher secondary exam scores scaled between1….3 ('1' for scores below 40%, ''2' for score-range 40%-60%, and '3 for score-range between 60%-100%), these scores were obtained from the examinations held for students, nationwide. The second addition was medium of instruction in either Hindi or English. The medium of study pursued by the students was mapped as: '1' for English medium and '2' for Hindi medium and the third was the living location which meant the region where student resides. The student's living location was scaled as: Rural-1, Sub-urban -2, Urban -3. These third set of experiments in the series was performed in order to find the impact on the precedence ordering of both static and dynamic attributes when considered in total.

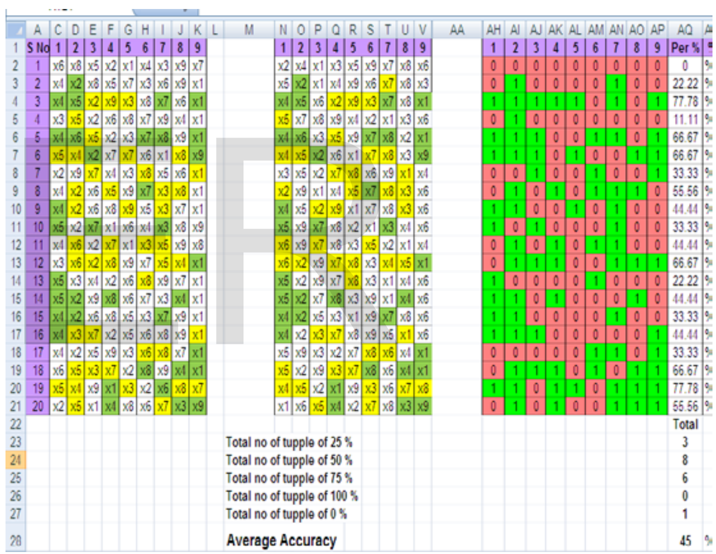

Fig 5.4 Attribute Precedence comparisons for Nine-Attribute Bayesian Driven FE Model Vs RELIEF (Prior-probabilities) FE Model, (Method II)

On repeating similar experiments as done for six-attribute Baysian driven FE model, the model accuracies of Nine-Attribute Baysian driven FE model were obtained, when compared with both the methods of RELIEF FE model. Such accuracy computations were summarized as tabulation shown in table 5.1 for all the three experimental modules.

TABLE 5.1 PERFORMANCE COMPARISONS OF FOUR, SIX, NINE-ATTRIBUTE PRECEDENCE RELATIONS OF FITNESS(DATASET I)

Model Type	**Comparison with RELIEF weights (normalized)**	**Comparison with RELIEF weights (prior-probabilities)**
Experiment_ID	**Accuracy (%)**	**Accuracy (%)**

4-ATTRIBUTE FE MODEL	82%	83%
6-ATTRIBUTE FE MODEL	60%	58.33%
9-ATTRIBUTE FE MODEL	39.44%	45%

5.3 Feature Extraction cum Ranking: Performance Analysis

In order to validate the overall significance of initial set of four attributes, entirely referring to students' academic efforts (put to Four-Attribute FE modeling experiment), experiments were conducted in increasing order of 'p' values; p denotes number of experimental attribute dimensions. The tabular observations show that for p=4, the model accuracy value ranges from 82% to 83%; for p=6, the accuracy reduces to span of 58 - 60% and for p=9, model accuracy drops down to 39 - 45%.

The performance comparisons conducted between innovative (Naïve Bayesian driven FE) model and benchmark methodology (RELIEF based FE model) can be visualized from accuracy figures of ranking attributes in increasing order of their relevance as compared to smaller accuracies obtained for larger spectrum of attributes with a mix of explicit and implicit ones

In order to confirm upon the authenticity and efficiency of the novel hybrid FE model upon up experimental data sets from the one study center, the experiments were repeated with real-time datasets with different academic settings. This time students' live data sets were collected from another academic institution for similar attribute schemas, but for different academic course domain viz. bachelors' degree of engineering, comprising 145 training students' instances and 50 test instances

TABLE5. 2 PERFORMANCE COMPARISONS OF FOUR, SIX, NINE-ATTRIBUTE PRECEDENCE RELATIONS OF FITNESS(DATASETII)

Experiment_ID	**Accuracy (%)**	**Accuracy (%)**
Model Type	**Comparison with RELIEF weights (normalized)**	**Comparison with RELIEF weights (prior-probabilities)**
4-ATTRIBUTE FEMODEL	83%	72.5%
6-ATTRIBUTE FEMODEL	62%	56.3%
9-ATTRIBUTE FEMODEL	39.78%	44.44%

The performance evaluation experiments showed similar patterns of model accuracies as that exhibited with students' datasets from previous academic course settings i.e. decreasing order of Feature Extraction FE model accuracies for increasing order of attribute schemas (p=4, 6 and 9) respectively as tabulated in Table 3. This ensured the generic functionality of the newly implemented Naïve. Bayesian driven F E model in ranking any number and any type of attributes with any kind of real-time academic domains.

As can be viewed from table 5.2, the FE model accuracies decrease with increasing dimensionality of attribute schema (feature vector length). It can also be inferred that the dynamic attribute reflecting the academic efforts of the students have a great impact on the validity of the Attribute Precedence Relations, as compared to their state attributes

6 CONCLUSION

The mentioned piece of research work open up the novel mining objectives in teaching-learning environments. It provides a precise description of a "stitch-in-time" methodology to be used to analyze and evaluate the rank order of academic attributes contributing to students' academic performance for their ongoing courses, well before they face their final (End Semester) examinations. The implemented framework came out with many types of valuable decision-making tips that shall help the management to take needful pre-emptive actions. Unlike the recent research trends in EDM that focus on promoting overall students' academic performance during their course tenures, the study is capable of providing student-wise precise recommendations towards remedial actions to be adopted by them (preferably slow learners). In this way, the students' grasping levels can be identified individually, preferably those suffering from weak academic profiles, may be due to varied underlying factors. These identified weak students can be counseled in varied directions (with necessary remedial actions) some of which are described as follows:-

- **Vocational Courses**: The students lacking skills to complete assignments in time can be encouraged to solve questionnaires supported by teachers' interaction that shall boost individual attention for enhancing their writing skills on presenting answers, although they may possess consistent level of knowledge.
- **Extra Classes**: The students found weak in specific subjects can be called for remedial classes owing to variety of genuine reasons for lagging behind in these subjects.
- **Tutorial Sessions:** This measure can be another remedial action for those who haven't appeared in internal assessment of all subjects or have appeared with every few subjects and showed poor performance. These sessions enable them to submit assignments, gaining confidence while solving past question papers and improving their writing skills too.
- **Usage of digital instructional aids**: At times, soft-

copy resources available from Web or from Interactive-Communication Technology promoted world-wide, help the students possessing average or less grasping skills for understanding concepts better than perceiving the same from blackboard teaching methodology.

- **Extra Laboratory Effort:** In order to achieve higher attainment in programming/ laboratory based theory subjects, extra laboratory sessions / programming practice may be required for some students to enhance the application-oriented reasoning / real-time simulation / logic formulation skills.

Adding to the interest of university, management and curriculum committees, and the recommendation-cum-ranking results can be exploited to arrive at assimilated figures of statistical measures in order to make amendments in the structure of courses and to encourage quality data collection processes.

The mentioned recommendation-cum-ranking model ensures the sustainable academic progress for all the participating EDM stakeholders (management, faculties, students and parents).

ACKNOWLEDGMENT

The authors acknowledge the resource centre for providing infrastructure support and live datasets to carry the mentioned work as a part of doctoral course in the discipline of Computer Applications

REFERENCES

[1] B. M. Bidgoli and W.F. Punch, "Predicting Student Performance: An Application of Data Mining Methods with the educational Web Based System LON_CAPA", G. A. Research and Application, Boulder, Nov 5-8, 2003.

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

[3] A. Qasim, AI-Radaideh and E. Al. Negi, "Using Data Mining Techniques to Build a Classification Model to Predicting Employees Performance", Int. J. Advanced C omputer Science and Applications, vol. 3, no. 2, 2012.

[4] M. Fire, G. Katz, Y. Elovici, B. Shapira and L Rokach, "Predicting Student Exam's Scores by Analyzing Social Network Data", Telecom Innovation laboratories and information Engineering Department, Ben-Gurion,Israel, Dec. 2012.

[5] R. Kohavi, H. John and Pfleger Karl, "Irrelevant Feature and the subset selection problem", In Machine Learning W.Cohen and HaymHiesh, eds. San-Francisco, CA, [Proceedings 11th Int. Conf. Isreal, pp. 121-129, 1994. .

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

[7] K. Kira and L.A. Rendell, "A practical approach to feature selection," Proc. 9th Int. Conf. Mach. Learn, pp. 249, 1992.

[8] Singh, M. and Rawal, A.(2010),Prediction of subject scores in ongoing courses for Academic Performance Using Decision Tree Techniques, [Proc. Nat. Conf. Information and Communication Technology Nagpur, India,

[9] Singh, M. and Singh, J. (2013), Machine Learning Technique for prediction of subject scores: A Comparative Study, International Journal of Computer Science and Network, Vol. 2, no.3, 2277-5420

[10] Singh, M., Singh, J. and A. Rawal.(2014),Int. Conf. on Information Technology,Silicon Institute of Technology Bhubaneswar,Orissa Information Soc., Bhubaneswar, 20.,

[11] Singh, M., Singh, J. and A. Rawal (2014), Performance Evaluation of Feature Extraction Model to Identify Student Appraisals, International Journal of Advanced Research in Computer and Communication Engineering Vol. 3 no.9, 7964-7968.,

[12] Singh, M., Singh, J. and A. Rawal. (2015), Student Academic Counseling from Attribute Precedence Relation using EDM, proc. Int. Conf. on Advances in Information Processing and Communication Technology, IPCT ,16-21..

[13] Y. Sun and D. Wu, "A Relief Based Feature Extraction Algorithm", Atlanta, Georgia, USA,[Proceedings Int. Conf. on Data Mining,,Apr. 2008.

[14] M. Sewell, "Feature Selection", 2007, http://machine-learning.martinsewell.com/feature-selection/feature-selection.pdfK

[15] A.A. Aziz, , N. H. Ismail, and F. Ahmad, "Mining students' academic performance," J. Theoretical and Applied Information Technology, vol. 53(3), pp. 485, July 2013.

[16] Ying Liu "A Comparative Study on Feature Selection Methods for Drug Discovery" Georgia Institute of Technology, College of Computing, Atlanta, Georgia 30322, April 13, 2004.