Applying Predictive Analytics in Elective Course Recommender System while Preserving Student Course Preferences

Ridima Verma
Computer Science & Engineering Department
TIET, Patiala, India
vermaridima@gmail.com

Anika Gupta
Computer Science & Engineering Department
TIET, Patiala, India
anikagupta2010@gmail.com

Abstract— **In higher educational scenarios, elective courses sought to provide a deeper insight of the trending advancements in the field of specialization for undergraduate students. So, choice of elective subjects during the pre-final or final year of the undergraduates play a crucial role as they help in shaping their career or area of specialization for future research. However, there exist numerous gaps and concerns that arise due to mismatch of the elective courses pre-requisites and the student's possessed skills-set which result in degraded quality as well as student academic performance. This research study focuses on filling in these gaps by predicting the marks in different elective subjects for the current cohort of students, beforehand, as well as side by side preserving their explicit subject preferences. With the help of the proposed methodology an accuracy of 88% was achieved for providing efficient bilateral elective course recommendations.**

Keywords—Educational Data Mining, Recommender Systems, Higher Education, Decision Tree, Regression, Support Vector Machine, Neural Network.

I. INTRODUCTION

In case of higher education, as the undergraduates proceed in their academics, they develop a liking for a particular field of study in their area of specialization. The higher academic course curriculum is also arranged in such a way that after imparting the knowledge of core subjects during the initial two and half years of their study, in pre-final or final year they are sought free to choose their own area of specialization in the form of different elective subjects. The elective subjects provide an insight into the trending advancements of their fields and form the basis of their career, as these tend to give a glimpse of their future area of research and specialization.

However, as per the current educational scenarios, the undergraduates remain mostly confused on what to choose as they either lack in having the sufficient initial knowledge of the elective subjects or are having knowledge overflow of all subjects and so are unable to decide which one to choose. In such scenarios, they often seek the advice of their instructors or friends and mostly go with the cohort choice. However, going with the flow often creates a gap between their actual skills set and the required skills set for the elective subject that they have preferred as their choice. In later stages, this results in loss of interest of the students in the enrolled elective subject and hence a degraded academic performance is encountered by the institution. On the other hand, from the institutional perspective also elective subjects play a vital role as they need to arrange all the required infrastructural and teaching resources timely for the successful run of the various elective subjects. In worst cases, institution is bound to choose another subject over the students preferred subject due to lack of the resources. Similarly, as a result of this, there can be numerous limitations, gaps or concerns arising either in case of students or institutions in real world educational scenarios.

This research study tries to bridge these gaps by recommending efficient elective course subjects to the institution that indirectly predicts the academic success rate of different elective courses beforehand and along with this also preserves the student subject explicit preferences. This approach can help to reduce the skill-gap often seen in case of higher education scenarios and timely guide the institutions for the electives that can result in overall increased academic performance and quality education. Figure 1 shows the proposed framework for generating such elective course recommendations.

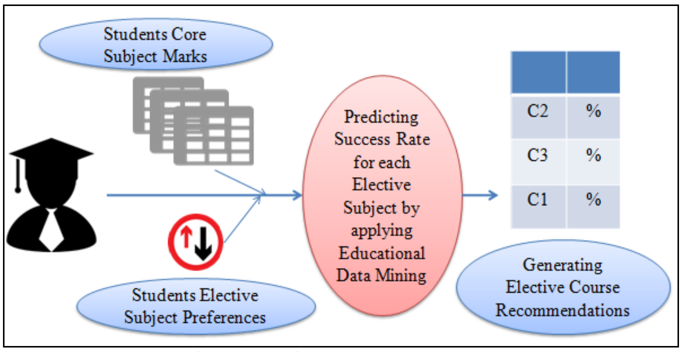

Figure 1: Proposed framework for Elective Course Recommender System

II. LITERATURE SURVEY

A lot of research has been done in the field of prediction of student academic performance and providing course recommendations. Apart from the number of features that influence student's academic performance, choice of models used for making the predictions also play a crucial role in the prediction procedure. A brief review of the various existing models or techniques is discussed below.

Sharma and Sharma [1] compared two collaborative filtering approaches for predicting grades of students based on their performance in earlier courses. Arsad *et. al.* [2] presented a neural network prediction model for predicting the academic performance of students based on multi entry level. Shahiri *et. al.* [3] presented a meta-analysis based on prediction methods to understand factors influencing students' academic performance. Romero *et. al.* [4] compared classification models for classifying students based on their final marks obtained in moodle courses. Osmanbegović and Suljić [5] evaluated learning methods for predicting success of students in courses, based on grades obtained by students in the considered courses. Bunkar *et. al.* [6] compared various classification models for predicting student performance in order to predict the final grade in a course under study. Guo [7] presented a mixed neural network and statistical prediction approach, which helped in establishing dynamic models for analyzing and predicting student's course satisfaction. Ramesh *et. al.* [8] applied techniques to predict student performance based on influencing factors along with their grades obtained in courses. Affendy *et. al.* [9] applied prediction models to obtain influencing factors in predicting student performance based on marks obtained by students in early stages of their course. Chamillard [10] applied predictive models to predict student performance based on percentage of marks and grades obtained by students in previous courses. Al-Badarenah and Alsakran [11] presented a collaborative filtering recommender system using clustering techniques for generation of elective courses association rules to recommend courses based on similarity measure. Cakmak [12] presented a collaborative filtering based enhanced model, incorporating outlier elimination and GPA based similarity for predicting students' grades in future courses. Tran *et. al.* [13] proposed a hybrid model by combining collaborative filtering and regression strategies for performance prediction of students. Mueen *et. al.* [14] proposed a model using data mining techniques to predict and analyze students' academic performance. Bydžovská [15] also proposed a prediction model using classification, regression and collaborative filtering approaches for predicting final grades of students based on previous achievements of similar students.

Despite of vast literature existing on the efficient models and methodologies being used for making efficient student academic course prediction, elective course recommendations area seems to be unexplored by the researchers. Not enough literature have been found that provides elective course recommendations to the institutions based on the student preferences as well as also merits over the estimated academic performance of the current batch. This study is aimed to bridge this gap and provides a bilateral elective course recommendation that assures student as well as institution success.

III. Aim and Scope Of The Study

For blended learning environments, course recommendations form a bilateral process as institutional resources and student interests both needs to be taken care off. In blended educational scenarios, institutional infrastructural and teaching resources often form a limiting factor. As a result, despite of the trending student subject preferences, institution based pre-defined subject list is imposed on the students considering the institution resource limitations at the last hour. On the student part, as per the student's disinterest in the imposed subjects, the academic performance of the student diminishes that result in quality degradation of the skill set possessed by the students. This indirectly also effects the learning outcomes of the subjects with lower success rates and thus affecting institutional quality education.

On the other hand, a similar case of considering only the student elective subject preferences and allotting them the desired subjects also does not assure complete academic success as there exists a gap between student's desired subject and the possessed skill-set. For practical scenarios both must go hand-in hand as the possessed skill-set or capability must match the course demanding pre-requisite and ongoing skills for better academic success.

Considering the present blended educational scenarios, there is a need of an elective course recommender system that is aimed to provide timely elective course recommendations for institutions that assures academic success of the students at the backend while preserving student subject interests. These predictions in turn, can also help to intimate the institution for the necessary arrangements of the resources, beforehand, that will lead to successful academic quality performances. So, the aim of this study is propose a methodology that can be implemented for generating efficient elective course recommendations for assuring two-sided success. This study targets the following research objectives:

1. Identifying the efficient data mining technique for predicting marks in the proposed elective subjects.
2. Propose an algorithm that also considers the contextual information of student varying preferences.
3. Provide efficient list of elective course recommendations that assures bilateral academic success of students as well as the institution.

IV. Methodology

This section discusses about the methodology adopted for generating elective course recommendations. The complete methodology flowchart is briefly described in Figure 2.

A. Dataset Description

Real time university anonymous dataset, for past 2 years, consisting of undergraduate student's core subject marks, subject preferences, student allocated elective subject and marks in the respective elective subject of computer science and engineering department during their final year, were used in this study. The dataset consisted of approximately 658 student entries with 13 attributes, namely marks obtained by students in their ten core subjects of computer science and engineering curriculum, actual allotted elective subject, actual marks in the elective allotted subject, along with student preferred interest subjects.

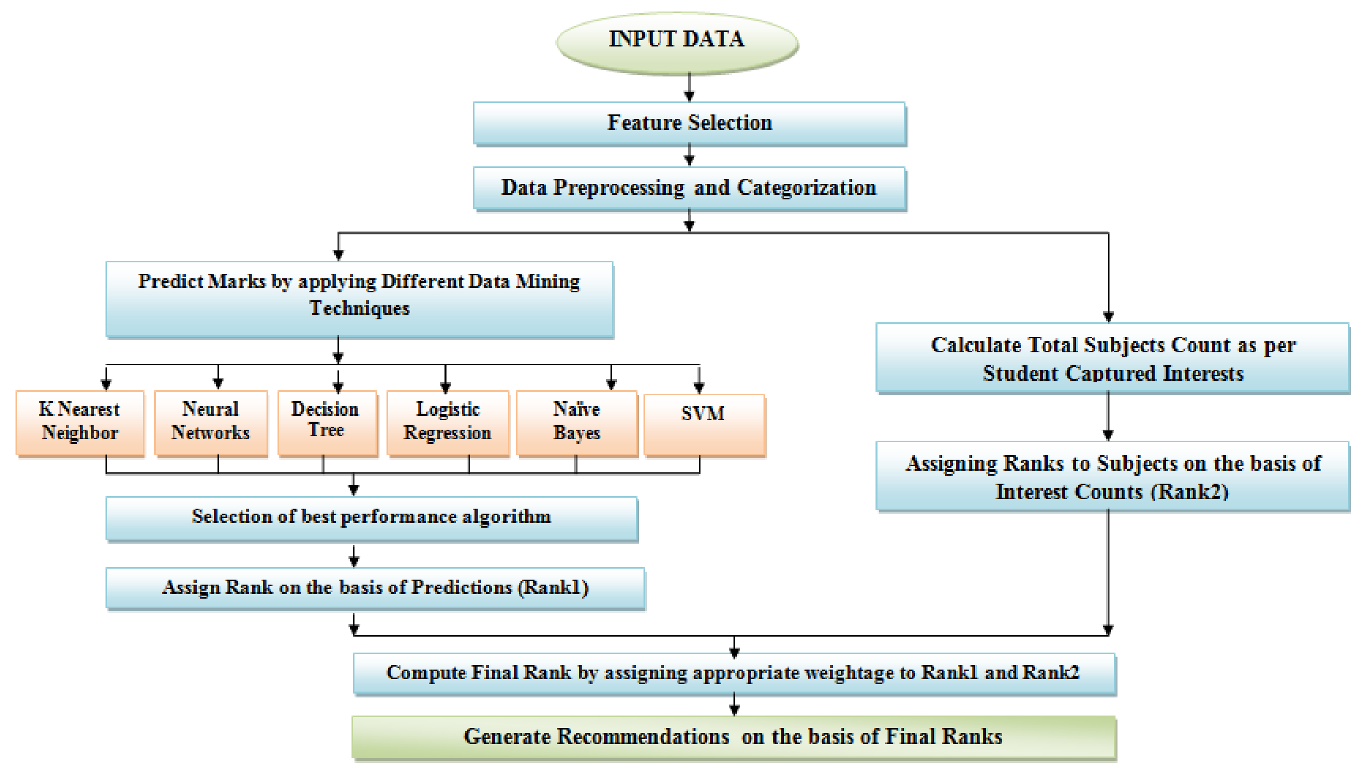

Figure 2: Complete flowchart of the proposed system for generating elective course recommendations

TABLE I: ELECTIVE COURSE SUBJECTS LIST

Code Used	Attributes
SVV	Software Validation and Verification
CC	Cloud Computing
CS	Cyber Security
IoT	Internet of Things
ML	Machine Learning
NLP	Natural Language Processing
AR_VR	Augmented Reality and Virtual Reality

TABLE II: SELECTED FEATURES LIST OF THE DATASET

No.	Attributes	Description	Selected ?
1	Student Gender	Male, Female	No
2	Student Family Status	Lower, Middle, Upper	No
3	Student Marks in 10 Core Computer Science and Engineering Subjects (Computer programming, Data Structures and Algorithms, Database Management Systems, Microprocessor, Computer Architecture, Software Engineering, Computer Networks, Operating System, Web Development, Automata)	Marks ranging from 0 to 100	Yes
4	Student Preference (Interest) for Elective Subject	Only first subject preference was considered	Yes
5	Student Allotted Elective Subject	Assigned elective subject	Yes
6	Student Marks in Allotted Subject	Marks ranging from 0 to 100	Yes

B. Feature Selection

Within the anonymous datasets, out of 15, 13 relevant features for the respective domain were selected. These selected features were further used for model construction and to train and test the models for obtaining efficient elective course recommendations. The 13 considered features consisted of 11 continuous attributes and 2 discrete attributes. Table I and II represents the considered elective course subjects list and a brief description about the actual features considered for this study, respectively.

C. Data Preprocessing and Categorization

As discussed, in the dataset out of thirteen, eleven were continuous marks attributes. Although, in small number, but the missing values in case of marks were replaced by the average value obtained by students for that particular subject. Missing values for allotted and preferred subjects attributes were not encountered as these were mandatory fields for every student.

Data categorization was also done within the dataset by converting the obtained marks in the allotted subject from continuous variable to a discrete variable through the process of binning of data. The allotted subject marks were binned into different classes or bins with a fixed difference of 5 marks each. So, the actual binning of data for allocated marks started from [0 to 5], [6 to 10], [11 to 15] and so on till [96 to 100]. The motive behind this categorization was to fit the data for various data mining classification techniques used further in this study. A glimpse of final pre-processed and categorized marks dataset is shown in Table III.

TABLE III: DATASET AFTER PREPROCESSING AND CATEGORIZATION PHASE

Marks S1	Marks S2	Marks S3	Marks S4	Marks S5	Marks S6	Marks S7	Marks S8	Marks S9	Marks S10	Allocated Elective Marks	Categorized Elective Marks
72	75	92	67	55	70	79	75	68	81	86	[86 to 90]
41	38	53	68	64	56	71	46	81	21	64	[60 to 65]

D. Applying Educational Data Mining

After data preprocessing and data categorization, various supervised learning classification models were imposed on the final dataset. The selection of the various classification techniques used for generating predictions were done on the basis of previous studies. According to the previous studies, six common classification models were identified that are frequently used by researchers and outperformed the marks based predictions on their respective datasets. The identified six techniques were: K-Nearest Neighbor used within collaborative filtering approach, Naïve Bayes, Decision Tree, Neural Network, Logistic or Linear Regression (for continuous data) that is mostly preferred for marks prediction and Support Vector Machines.

So, further a brief description of the final six models (K-Nearest Neighbor, Neural Network, Decision Tree, Naïve Bayes, Support Vector Machines and Logistic Regression) applied on the final datasets is given below.

k-Nearest Neighbor

k-Nearest Neighbor (kNN) forms the basis of the collaborative filtering technique used in recommender systems. kNN is a simple algorithm which classifies new cases based on similarity measure. It is based on feature similarity. This is better choice for classifications where accuracy is important. For predicting the various categories of allocated subject marks various values of k (1, 2, 5, 10, 20,…, 100) were experimented. Initially the accuracy varied on larger scales but with the k-size between 50-60, the results yielded highest accuracy and after this the accuracy was stabilized. So, the experimentation proceeded by assigning a k-size of 60.

Neural Network

Neural Network (NN) simulates the principle of human brain that is composed of highly inter-connected network of neurons. The data nodes in this model act as input nodes of the neurons. The model is trained by iteratively adjusting the weights of the hidden layers in order to obtain the desired output. Neural Network thus modifies itself as it progresses further and learns from the training sets. However, neural network is said to work faster on large datasets. Here, the model was trained by supplying of ten core subject marks as input data nodes and actual allotted subject marks category as desired output.

Decision Tree

The Decision tree (DT), used in supervised learning, are tree like structures, often used as decision support tools. Decision trees contains conditional control statements and promise an output in either case *i.e.*, whether the condition is met or not. This model breaks down the dataset into smaller subsets on the basis of conditions and also incrementally develops the tree along side. The result of this is a decision tree with decision nodes as conditions and leaf nodes as outputs. For the datasets used, decision tree's ID3 algorithm was used for classification and allotted subject marks category as output. Figure 3 shows a small portion of the decision tree generated for the dataset.

Naïve Bayes

Naïve Bayes (NB) classifiers are probabilistic classifiers based on Bayes theorem and assume that the features involved are of independent nature. It uses maximum likelihood for parameter estimation. Because of the excessive use of Naïve Bayes classifier and its variants in educational scenarios for student knowledge estimation, this was also selected and used for making the prediction for the allocated subject marks category of the students.

Support Vector Machine

In supervised learning, Support Vector Machine (SVM) is used as a discriminative classifier which classifies on the basis of separating hyper-plane. When it is supplied with supervised learning training dataset, it outputs an optimal hyper-plane that further helps in classifying or categorizing the testing data. This forms the first choice when the dataset size is small. Inputting the labeled dataset, the classifier was trained and further used to test and predict the allocated subject marks category as the output.

Logistic Regression

Linear Regression is most commonly used for marks prediction. But due to the categorical nature of the target variable here, Logistic Regression was used for predictive analytics. Logistic Regression analysis is similar to linear regression except that the outcome in case of Logistic Regression is dichotomous. However, because of its flexibility and adaptive nature it can be used for categorical data too. The categorical variable considered for prediction was the allocated marks subject attribute.

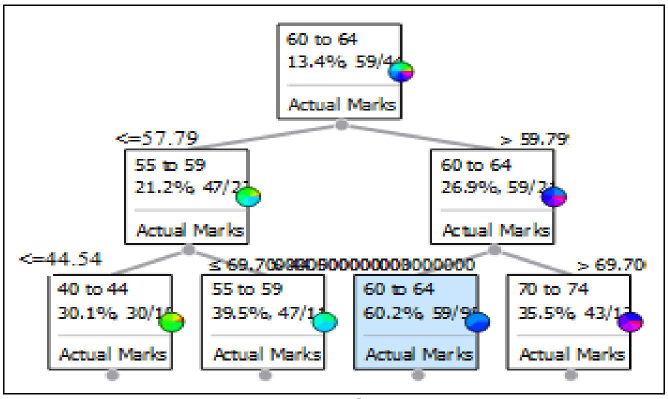

Figure 3: Generated Decision Tree

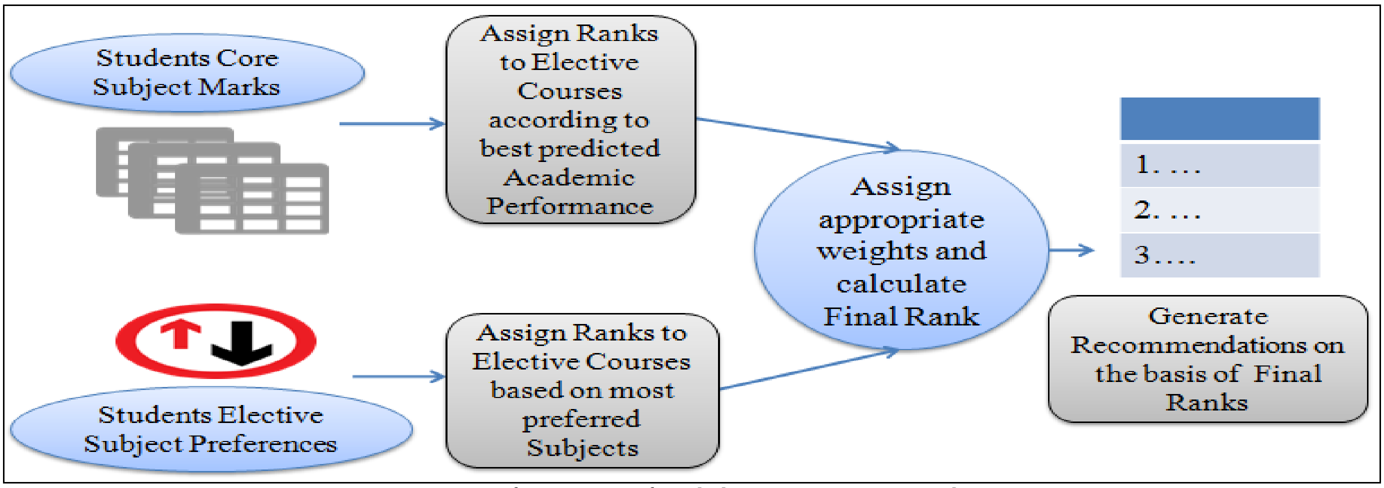

Figure 4: Steps for generation of Final Elective Course Recommendations

E. Rank-I Calculation

After the application and evaluation of the different data mining techniques used in this study, the model producing the highest accuracy was selected. In this case Support Vector Machine's were used as it was found to be the best in terms of accuracy when comparing different data mining techniques used in this study. After this the marks predictions were obtained for each elective separately. Once the predicted marks categories for each elective subject for each student were obtained, efforts were put for finding the average marks category slots for that entire elective subject using the student allocated subject column. Cumulative frequency concept was used for determining the average marks category slot for the particular subject. Once the average marks category slot were calculated for each elective subject, ranks were assigned to the subjects, where highest average marks category was awarded with rank 1 and lowest average marks category with rank n, n being the total number of elective subjects (7 in this case).

F. Rank-II Calculation

In parallel, the student subject preferences were also considered and actual count of students preferring a particular subject was calculated. This procedure was repeated for each existing elective course. Once all the counts were calculated, again ranks were provided based on the most preferred and least preferred elective subject. Highest preferred subject was assigned a rank of 1 and least preferred a rank of n (n is 7 in this case).

G. Weighted Ranks Calculation

Since efforts are put in to preserve the student preferences along with assuring quality academic performance of the students collectively, the weighted rank concept was proposed. After the calculation of Rank1 and Rank2 from the above steps, final weighted rank can be calculated by assigning w1 to Rank1 and w2 to Rank2 as given in the following equation:

$$Weighted\ Rank_i = w1 * Rank1_i + w2 * Rank2_i \qquad (i)$$

Here,
$Rank1_i$ = Rank1 for i^{th} elective subject preserving academic performance for i^{th} elective subject.
$Rank2_i$ = Rank2 for i^{th} elective subject preserving student subject preference.
w1 = 0.5
w2 = 0.5
In this study, equal weightage was given to both, so w1 and w2 both were assigned equal weights of 0.5. However, these can further vary depending upon the institutional requirements and preferences.

H. Final Rank Generation

On the basis of weighted ranks obtained in earlier step, final ranks were calculated again, by assigning most preferred (having lowest weight) a rank of 1 and least preferred elective subject with a rank of n (n is 7 in this case).

TABLE IV: CALCULATION STEPS USED FOR GENERATING FINAL RANKS

Elective Subjects	Average Marks Category Slot	Rank1	Subject Preference Count	Rank2	Weighted Ranks	Final Ranks
Cloud Computing	[60-64]	4	50	7	(4*.5+7*.5)= 5.5	7
Internet of Things	[55-59]	5	68	2	(5*.5+2*.5)= 3.8	4
Natural Language Processing	[70-74]	2	55	6	(2*.5+6*.5)= 3.6	3
Cyber Security	[45-49]	7	65	3	(7*.5+3*.5)= 5.4	6
Software Verification and Validation	[65-69]	3	62	4	(3*.5+4*.5)= 3.4	2
Machine Learning	[75-79]	1	79	1	(1*.5+1*.5)= 1	1
Augmented and Virtual Reality	[50-54]	6	62	4	(6*.5+5*.5)= 5	5

The subjects are then arranged in increasing order of their Final Ranks and recommended to the concerned authorities for further considerations. Figure 4 and Table IV briefly describes the complete procedure of generating the elective course recommendations.

V. RESULTS AND DISCUSSIONS

For experimentation, educational data mining, analysis and predictions, python's Orange library was used [16]. The final results of the predictions were stored in the form of excel sheet which was further processed as per the proposed algorithm. The analysis was performed on a dataset consisting of 658 rows. Each data mining model used their own learning algorithm and generated a model that fitted the input data and generated predictions.

Predictions that are generated via models are expected to have good generalization capability and can efficiently produce a correct class label or categorization for the previously unknown data. Classification model's performance is evaluated on the basis of how many correct and incorrect predictions are made by the model on the testing dataset. The confusion matrix for a two-class classification problem is shown in Table V.

For efficient comparison of these different data mining models used within this study, four different evaluation metrics were used for judging the quality of the predictions made, namely: accuracy, precision, recall and F1 score. Also, for training and testing purpose of each data mining model, 10-fold cross validation approach was used for statistical analysis of the datasets.

A. Accuracy

Accuracy is the measure of degree of closeness of the quantity's measurement to that quantity's true value. It can be calculated from the confusion matrix using the following equation:

$$Accuracy = (TP + TN) / (TP + TN + FP + FN) \qquad (ii)$$

where TP, FP, TN, FN being number of True Positive, False Positive, True Negative and False Negative, respectively. The accuracy (in percentage) for different educational data mining models used, in case of different elective subjects is shown in Figure 5.

B. Precision

It is the ratio of correctly predicted positive observations to the total predicted positive observations. It can be calculated from the confusion matrix using the following equation:

$$Precision = TP / (TP+FP) \qquad (iii)$$

where TP and FP stands for True Positive and False Positive respectively. The precision value ranging from a scale of 0 to 1 for different data models used, for different elective subjects is shown in Figure 6.

TABLE V: CONFUSION MATRIX FOR BINARY CLASSES

		Predicted Class	
		Class = 1	**Class = 0**
Actual class	**Class = 1**	True Positive (TP)	False Negative (FN)
	Class = 0	False Positive (FP)	True Negative (TN)

C. Recall

Recall is defined as the number of true positives over the number of true positives plus the number of false negatives. It can be calculated from the confusion matrix using the following equation:

$$Recall = TP / (TP+FN) \qquad (iv)$$

With TP and FN stands for True Positive and False Negative respectively. The recall value ranging from a scale of 0 to 1 for different data models used, for different elective subjects is shown in Figure 7.

D. F1 Score

For binary classification, F1 score measures the test accuracy. It is the weighted harmonic mean of precision and recall. It is calculated by considering both recall and precision as given in the following equation:

$$F1\ Score = 2 / ((1/Recall)+(1/Precision)) \qquad (v)$$

The F1 score value ranging from a scale of 0 to 1 for different data models used, for different elective subjects is shown in Figure 8.

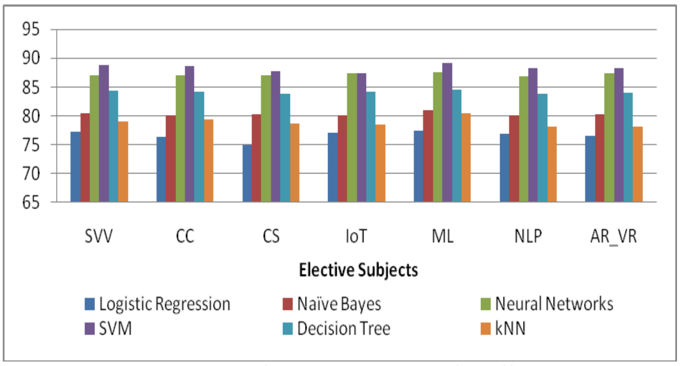

Figure 5: Comparison of accuracy percentage for different models

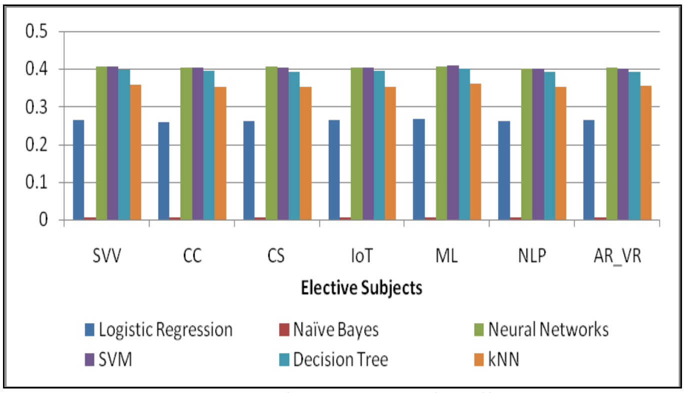

Figure 6: Comparison of precision value for different models

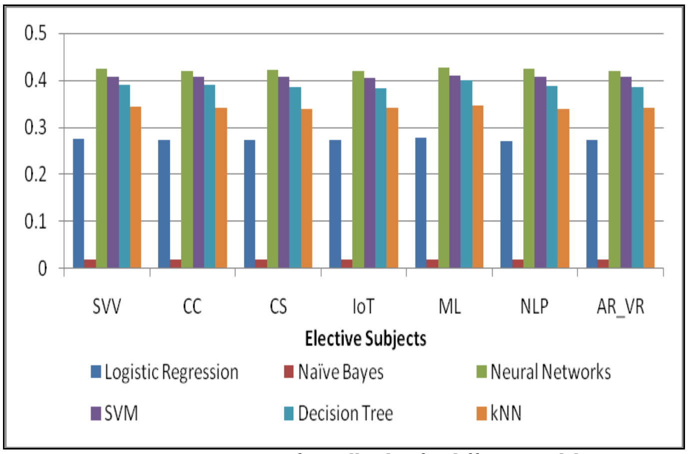

Figure 7: Comparison of Recall value for different models

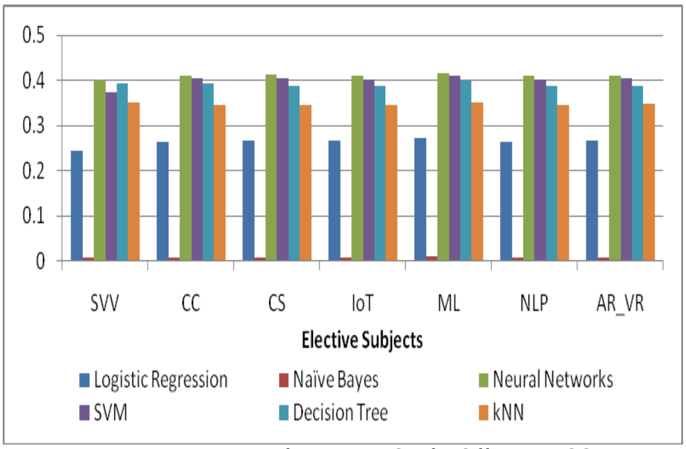

Figure 8: Comparison of F1 Score value for different models

After analyzing the different models used in this study on the performance evaluation parameters Support Vector Machines outperformed in terms of accuracy with an average accuracy of 88 percent across all the elective subjects. Support Vector Machines were ahead in terms of Accuracy whereas Neural Networks were ahead by fractions, in all other evaluation parameters, namely Precision, Recall and F1 Score. Except Support Vector Machines and Neural Networks that yielded somewhat comparable results, other models were far behind in terms of accuracy, precision, recall and F1 score. An exceptional behavior of Naïve Bayes which have exceptionally low values in precision, recall and F1 score parameters is also observed.

Naïve Bayes specifically focuses on the independence of the features and hence, this property of Naïve Bayes can self-explain the reason behind the lower values of precision, recall and F1 score.

Out of Neural Network and Support Vector Machines, Neural Network is said to have higher accuracy, but as already discussed, Neural Network requires large amount of data for the efficient training of its hidden layer nodes. However, here as the dataset considered is small, so this can be the reason why Support Vector Machines produced higher accuracy as compared to the Neural Networks, because Support Vector Machines act as extremely good classifiers on smaller datasets. Support Vector Machines with the help of its hyper-planes, yielded an accuracy of 88.5 percent in one of the elective subject datasets. However, considering other evaluation parameters of precision, recall and F1 score, Neural Network model can be said to have better adapted for the considered dataset and in case of large educational datasets it may in future, even yield better and promising results. A comparison of the proposed system with other similar or partially similar systems and approaches is shown in Table VI.

VI. LIMITATIONS

This study is domain restricted, as the datasets used were from a single department. For better generalization, one can have comparatively larger cross-domain datasets. Also having less than 50 percent values in case of precision and recall, may account for the fact that the core subjects marks considered for the study were ten in number and must have lead to a large feature space. And training a large feature space requires sufficient larger datasets. But since the educational datasets having large amount of data and satisfying one's research requirements are not easily available and in direct university scenarios there remains security issues and educational privacy policies violation. So, the data collected here was taken on the consent of the students and all the information fields acquired were first anonymized for assuring the privacy of the students.

Another limitation of this study is that, here each subject marks were treated as independent observations. However, in real world scenarios, the marks obtained by students in different subjects are correlated at the backend. For instance, these can be correlated on the basis of the nature of the subject. For example, a subject can have a theoretical or a practical background. And the student enrolled in the subject can have a good hands-on practical experience. So, he may have higher marks in all those subjects that involve practical aspects as compared to those subjects that carry a theoretical background. This fact was not leveraged upon in this study as all marks were treated as independent observations due to anonymization of the datasets.

VII. CONCLUSION

In this study efforts were put to propose an efficient algorithm that assures academic success in the elective courses via its predictions as well as preserves the student subject preferences for greater achievement of bilateral academic quality learning outcomes. The distinction of this approach lies in the fact that it assures two-way success and takes care of institutional and student preferences at the same time. While considering the academic predictions, support vector machines were found to be the best predictor classifier model for predicting the marks in the respective elective subjects based on past academic score obtained in core subjects. Once the student subject preferences and subjects having highest academic success rate are identified, with the help of weighted ranks, the proposed algorithm helps to generate efficient elective course recommendations that can be used for assuring bilateral academic success of students as well as the institutions.

TABLE VI: COMPARISON OF PROPOSED SYSTEM APPROACH WITH OTHER EXISTING APPROACHES

REFERENCE	DATASET	TECHNIQUES/ALGORITHMS	RESULTS	LIMITATIONS
[1] Sharma & Sharma	Real time data (255 students)	User-based and Item-based Collaborative Filtering	Predicted grades of students. (MAE=0.38)	Smaller Dataset Single Model Approach
[2] Arsad & Buniyamin	Real time data (391 matriculation and 505 diploma students)	Neural Network	Predicting students' academic performance. (MSE=0.0409 for matriculation students and MSE=0.0488 for diploma students)	Smaller Dataset Single Model Approach
[4] Romero *et. al.*	Real time data (438 students)	Decision tree, Rule induction algorithm, Neural Network, Statistical Classifier, Fuzzy Rule	Predicted students' final marks. (Decision Tree outperformed with Accuracy= 67.02%)	Smaller Dataset Less Attributes
[5] Osmanbegović & Suljić	Real time data (257 students)	Naïve Bayes, Multi Layer Perceptron (MLP), C4.5	Predicted success in a course. (Naïve Bayes outperforms with Accuracy= 76.65%)	Smaller Dataset Less distinction between attributes
[7] Guo	Real time data (43 course data)	Neural Networks	Identified most relevant factors in student courses (MSE=0.0016 and Correlation=0.943)	Very Small Dataset Single Model Approach
[8] Ramesh *et. al.*	Real time data (900 students.)	Naïve Bayes, Multilayer Perception (MLP), SMO, J48, REPTree	Predicted grades of students. (MLP outperformed with Accuracy=72.38%)	Smaller Dataset Results based on psychometric factors
[10] Chamillard	Real time data. (285 students)	Regression Analysis	Predicted student performance. (Correlation=0.579)	Smaller Dataset Single Model Approach
[11] Al-Badarenah *et. al.*	Real time data. (2000 students)	Collaborative filtering, K-Means, Similarity distance measure.	Recommended elective courses. (Precision=0.95, Recall=0.5)	More algorithms can be used
[12] Cakmak	Real time data. (55475 course data)	User–based collaborative filtering, Item–based collaborative filtering, Similarity distance measure.	Predicted students' grades for recommending new elective courses. (MAE=0.32)	Single Model Approach
[14] Mueen *et. al.*	Real time data. (60 students)	Naïve Bayes, Neural Network, Decision tree	Predicted student academic performance. (Naïve Bayes outperformed with Accuracy=86%)	Very Small Dataset
Proposed System	Real Time data (658 students)	Naïve Bayes, Neural Network, Decision tree, Logistic Regression, Support Vector Machines, k-Nearest Neighbor	Elective Course Recommendations by preserving student interests (Support Vector Machines outperformed with Accuracy=88%)	Smaller Dataset Subject marks treated as independent observations

This research could be further enhanced as its future work, by considering cross-domain large datasets from different educational institutions for better generalized results. Other contextual attributes can also be researched upon for incorporating more efficiency within the proposed algorithm for generating more efficient elective course recommendations.

REFERENCES

[1] S. Ray and A. Sharma, "A collaborative filtering based approach for recommending elective courses," in *Springer International Conference on Information Intelligence Systems, Technology and Management*, 2011, pp. 330-339.

[2] P.M. Arsad and N. Buniyamin, "A neural network students' performance prediction model (NNSPPM)," in *IEEE International Conference on Smart Instrumentation, Measurement and Applications (ICSIMA)*, 2013, pp. 1-5.

[3] A.M. Shahiri and W. Husain, "A review on predicting student's performance using data mining techniques," *Procedia Computer Science*, vol. 72, pp. 414-422, 2015.

[4] C. Romero, S. Ventura, P.G. Espejo and C. Hervás, "Data mining algorithms to classify students," in *Educational Data Mining 2008*, 2008, pp. 1-10.

[5] E. Osmanbegović and M. Suljić, "Data mining approach for predicting student performance," *Economic Review*, vol. 10, no. 1, pp.3-12, 2012.

[6] K. Bunkar, U.K. Singh, B. Pandya and R. Bunkar, "Data mining: Prediction for performance improvement of graduate students using classification," in *IEEE Ninth International Conference on Wireless and Optical Communications Networks (WOCN)*, 2012, pp. 1-5.

[7] W.W. Guo, "Incorporating statistical and neural network approaches for student course satisfaction analysis and prediction," *Expert Systems with Applications*, vol. 37, no. 4, pp. 3358-3365, 2010.

[8] V. Ramesh, P. Parkavi and K. Ramar, "Predicting student performance: a statistical and data mining approach," *International Journal of Computer Applications*, vol. 63, no. 8, pp. 35-39, 2013.

[9] L.S. Affendey, I.H.M. Paris, N. Mustapha, M.N. Sulaiman and Z. Muda, "Ranking of influencing factors in predicting students' academic performance," *Information Technology Journal*, vol. 9, no. 4, pp. 832-837, 2010.

[10] A.T. Chamillard, "Using student performance predictions in a computer science curriculum," *ACM SIGCSE Bulletin*, vol. 38, no. 3, pp. 260-264, 2006.

[11] A. Al-Badarenah and A. Jamal, "An automated recommender system for course selection," *International Journal of Advanced Computer Science and Applications*, vol. 7, no. 3, pp.1166-175, 2016.

[12] A. Cakmak, "Predicting student success in courses via collaborative filtering," *International Journal of Intelligent Systems and Applications in Engineering*, vol. 5, no. 1, pp.10-17, 2017.

[13] T. Tran, H. Dang,V. Dinh,T. Truong,T. Vuong and X.Phan, "Performance prediction for students: a multi-strategy approach," *Cybernetics and Information Technologies*, vol. 17, no. 2, pp. 164-182, 2017.

[14] A. Mueen, B. Zafar and U. Manzoor, "Modeling and predicting students' academic performance using data mining techniques," *International Journal of Modern Education and Computer Science*, vol. 8, no. 11, pp. 36, 2016.

[15] H. Bydžovská, "A comparative analysis of techniques for predicting student performance," in *Proceedings of the 9th International Conference on Educational Data Mining*, 2016, pp. 306-311.

[16] J. Demsar, T. Curk , A. Erjavec , M. Milutinovicˇ, and M. Mozina , "Orange: Data Mining Toolbox in Python," *The Journal of Machine Learning Research* , vol. 14, pp. 2349–2353, 2013.