**COURSE RECOMMENDATION SYSTEM TO IMPROVE LEARNING RATE OF STUDENTS IN THE COVID SCENARIO**

**MINOR PROJECT REPORT**



Group Members:

Teghdeep Kapoor

Vardhika Jain

Tanya Pandhi

**ABSTRACT**

In higher education, Courses ought to provide a deeper insight of the trending advancements in the field of specialization for undergraduate students. Making choice of elective courses during the pre-final or final year of the undergraduates play a crucial role as it helps in shaping their career or area of specialization for the better learning. 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. 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. A personalized recommender system recommends efficient course subjects to the students that indirectly predicts the academic success of different courses beforehand and along with this also preserves the student subject interests.

**INTRODUCTION**

**E-Learning Recommender Systems**

The main task of these systems is to generate learning task recommendations based on learner’s previous tasks and performances. The likeliness of the learner can be constructed using the user profiles, which identify the needs of user. Recommendation generation task is broadly divided in two major phases: “learning” and “advising” phase. The former phase identifies learning patterns of previous users and the later phase helps in application of learned model for generation of recommendations.

Recommender systems can facilitate a user to find related items according to their interest. The main objective of these systems is to make selected recommendations based on user requirements. They implement the recommendation generation process by asking the user to rate a series of objects, on basis of which new recommendations are generated for the user and other similar users. These evaluations act as input for model generation, which helps in making predictions for the user on basis of either their profiles or the objects evaluated by them in the past.

The recommender systems must be able to produce tailored suggestions based on specific needs, which can help educators to know the strengths and weakness of learner and take measures according to learner’s skills. They should be able to facilitate a mechanism to compile the huge amounts of user data altogether at once to produce standard recommendations without any sort of discrepancy.

**Recommender System Approaches**

* Collaborative filtering (CF) -: It is a recommender system approach that is commonly used to generate personalized recommendations. It computes the likeliness either between clients or things. It leverages product transactions to produce recommendations. In this type of approach, for a specific customer, we find similar customers based on transaction history and recommend items that the customer in question hasn’t purchased yet and which the similar customers are likely to buy.
* Content Based filtering -: The main task of content-based filtering or cognitive filtering is to generate recommendations on basis of similarity between the list of the items and user profiles. The content of each item is pictured as a set of descriptors or terms, usually the words that occur in a document. Such systems leverage product information for its recommendations. For example, if a person looks at a book on an online bookstore, a content-based system would probably recommend the similar author books, because it would look at the author field.
* Knowledge Based Recommender-: The main task of knowledge-based recommender system is to use information about users and items and continue a knowledge based perspective for generation of recommendations, reasoning about what items meet user’s perspectives. Such recommender systems are more specific about the task performed by them based on explicit information about item assortment, user preferences and recommendations. These systems work well where approaches like collaborative filtering and content-based filtering cannot be applied. Non-existence of cold-start problem makes knowledge-based recommender system approach better than other existing approaches. For generation of explicit recommendations, there is a need for potential knowledge acquisition bottleneck to be triggered.
* Hybrid Recommender-: The combination of two or more recommender system approaches for better performance and lesser drawbacks constitute to formation of hybrid recommender systems. Most frequently, collaborative filtering approach is combined with other existing techniques in an attempt to avoid the problem of expansion. Combination of two or more techniques for building hybrid recommender systems seek to inherit advantages and disadvantages. Establishment of synergy between two or more approaches led to more accurate results for hybrid recommender systems that combine multiple recommendation techniques.

In spite of existence of different recommender system approaches, a number of approaches are practical to merge (i.e. Collaborative, Content-based and Knowledge-based Recommender), work will mainly focus on the combination of collaborative filtering and content based filtering techniques. Depending on the type of data and attributes, different types of combinations might produce dissimilar outputs.

**LITERATURE REVIEW**

Colossal amounts of research have been done in the field of prediction of student academic performance and provision of course recommendations. Apart from the number of features that can 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.

According to Funda Dag *et. al.*[1] approaches, some researchers have investigated that presentation of learning content and learning tools are designated based upon learning styles in the online learning environments is a factor that impacts the academic achievements of the learner. In the other research approach, researchers have used learning styles as a supportive factor to design the online learning environments for personalized online learning. As a result, it was seen that improving the academic achievements in online learning not only learning styles by itself are utilized on online learning and also the motivation of the learner, demographics factors, teaching strategies, and teaching methods should be considered.

According to Cathy Li *et. al.*[2], The impact of Covid-19 has dug up the roots of our education system. It has gone completely online, boosting the business of educational apps with higher chances of earning huge profits in the long run. Research suggests that online learning has been shown to increase retention of information, and take less time, meaning the changes coronavirus have caused might be here to stay.

According to S. Ray *et. al.*[3], In this paper they proposed a recommender system approach that made use of collaborative filtering approach for generation of elective course recommendations. They made use of both user-based and item based filtering, which is applied on real-time data for prediction of elective courses. Results are based on mean absolute error, calculated for each elective course.

# Jinjin Liang *et. al.* [4], proposed a hybrid teaching mode utilizing machine learning algorithms, which uses clustering analysis to analyze the learner’s characteristics and introduces a support vector machine to predict future learning performance. The hybrid mode matches the predicted results to carry out the offline teaching process.

H. Bydžovská [5], in this paper he proposed a prediction model using classification, regression and collaborative filtering approaches for predicting final grades of students based on previous achievements of similar students

D. Upendran *et. al.*  [6], they proposed a course recommendation system that undertook students as basis of their past performance and learning ability. They constructed model by using previous student data as input. The basic subsequent belief for the technique used is that if a student with certain skill is able to complete the course successfully then a new student with similar skills will be able to complete the course.

Arsad *et. al.* [7] proposed a neural network student performance prediction model which gave review on what factors can influence the final GPA obtained by students in their last semester. The results calculated were based on correlation and mean squared error between models applied.

Shahiri *et. al.* [8] proposed a recommender system for online course enrolment, which used historical data to show the factors which influences student’s choice of elective course selection. They made use of collaborative filtering approach and provided performance based on recall and coverage

Romero *et. al.* [9] proposed data mining techniques for classification of students which made structuring and implementation easier for instructors. They made use of pre-processing techniques such as discretization and rebalancing on collected data in order to obtain better classification results.

Osmanbegović and Suljić [10] presented data mining techniques for making student predictions based on their performance. Different data mining techniques were compared for generation of prediction model, which in return predicted student success. Data was collected by performing survey during the semester. The success of student was evaluated based on grades obtained by students in their final exam.

Bunkar *et. al.* [11] proposed data mining classification techniques. The main approach adopted by them was based on improvement of student performance. They made a 10 system that facilitates the use of rule generation process. Final grades of graduate students are predicted using decision tree model.

Guo [12] presented neural network approach by making use of statistical methods. These techniques incorporated establishment of dynamic models which helped in predicting student course satisfaction. Out of all the models applied, MLP outperforms in generation of near practical results.

Ramesh *et. al.* [13] presented statistical and data mining approaches. The objective of this study is to make student predictions based on their performance in the curriculum. They also considered influencing factors that can affect student performance and their final grades.

Affendy *et. al.* [14] proposed a model based on data mining approaches. They made predictions for calculating student academic performance based on factors that can affect their overall grades. The main objective of this study is to rank the affecting factors in order to warn the students, so that they are able to maintain their grades.

Chamillard [15] proposed a model based on data mining techniques. The main objective behind this study is to make an analysis in curriculum for generation of approaches that can make predictions based on each student’s past performance in their academic curriculum. This approach can help in making student performance predictions in their later courses.

Al-Badarenah *et. al.*[16] presented a collaborative filtering recommender system using clustering techniques for generation of elective courses by making use of association rules to recommend courses based on similarity measure.

Cakmak [17] proposed a recommender system approach that is used for estimation of student course grades. Along with this, they enhanced the proposed approach by implementing automated outlier removal, which improved the quality of result generation.

Tran *et. al.* [18] presented analysis on prediction of student performance in educational scenarios. They made use of different regression and data mining strategies for implementation of proposed approach. They also implemented combination of aforementioned techniques and calculated results based on model accuracies.

Mueen *et. al.* [19] presented an analysis for application of data mining techniques for prediction of student performance. They made use of real-time data collected from undergraduate students.

**PROBLEM STATEMENT**

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.

This research study tries to bridge gaps by recommending elective course subjects to institutes that indirectly predicts success rate of different elective courses beforehand and also preserves the student subject explicit preferences. These predictions in turn can also help to intimate the institution for the necessary arrangements of resources, beforehand, that will lead to successful academic quality performance.

Thus, the aim of this study is to propose a methodology that can be implemented for generating efficient elective course recommendations for assuring student as well as institution success.

.

The study targets the following research objectives:

* Identifying the efficient data mining techniques for predicting marks in proposed elective subjects.
* Propose an algorithm that also considers the contextual information of student varying preferences.
* Generate efficient list of elective course recommendations that assures bilateral academic success of students as well as the institutions.

**METHODOLOGY**

**Data Set Description**

Real time university anonymous dataset, for 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 information technology department during their pre-final year, were used in this study. The dataset consisted of approximately 2944 student entries with 13 attributes, namely marks obtained by students in their ten core subjects of computer science, actual allotted elective subject, actual marks in the elective allotted subject, along with student preferred interest subjects.

**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 1 and 2 represents the considered elective course subjects list and a brief description about the actual features considered for this study, respectively.

|  |  |
| --- | --- |
| **Code Used** | **Attributes** |
| 1. SVV | System Verification and Validation |
| 1. CC | Cloud Computing |
| 1. CS | Cyber Security |
| 1. IOT | Internet of Things |
| 1. AR VR | Augmented Reality and Virtual Reality |
| 1. ML | Machine Learning |
| 1. NLP | Natural Language Processing |

TABLE 1: ELECTIVE COURSE SUBJECT LIST

|  |  |  |
| --- | --- | --- |
| **Attributes** | **Description** | **Selected** |
| Student Gender | Male , Female | No |
| Student family status | Lower, middle or higher | No |
| 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 |
| Student Preference (Interest) for Elective Subject | Only first subject  preference was considered | Yes |
| Student Allotted Elective  Subject | Assigned elective  Subject | Yes |
| Student Marks in Allotted  Subject | Marks ranging  from 0 to 100 | Yes |

TABLE 2: SELECTED ATTRIBUTES LIST FROM THE DATASET

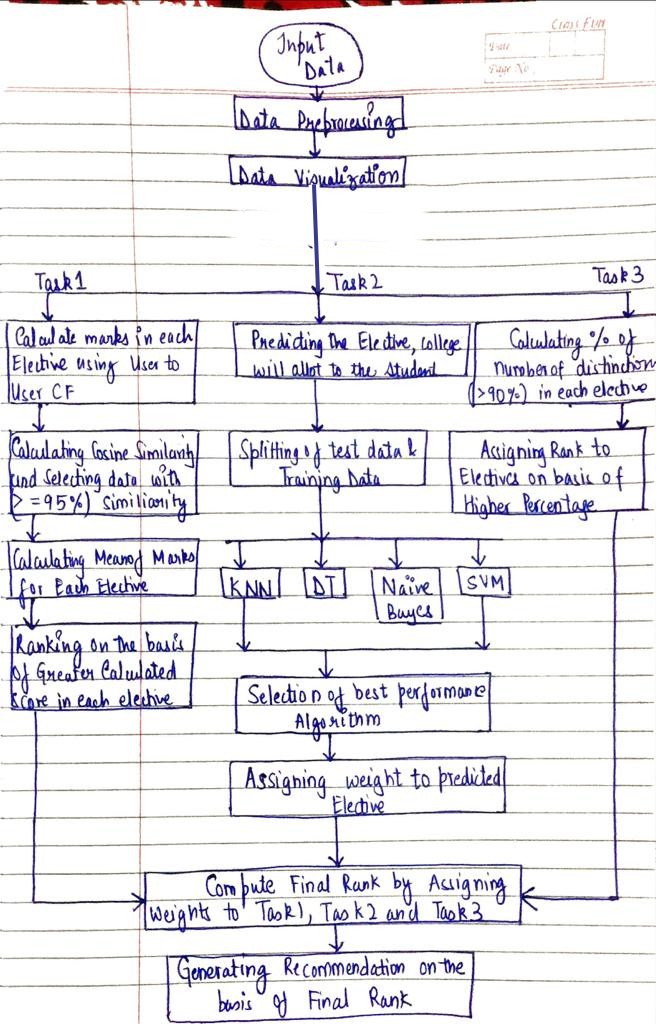
**Data Preprocessing**

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. Attribute Preprocessing was done within the dataset by converting all string literals in Preferred and Allotted Subject column to integer value.

Data was estimated to nearest multiple of five to reduce entropy/ randomness of the data.

The motive behind this preprocessing was to fit the data for various data mining classification techniques used further in this study.

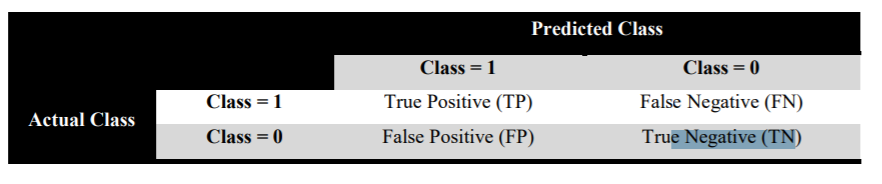
The complete methodology adopted for generating elective course recommendations is shown in Figure



Complete Flowchart of Proposed System for generating Elective Course Recommendations

**Applying Data Mining Techniques**

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 analyzing the elective allotment pattern of college on the basis of core subject marks and preference of elective. According to the previous studies, four common classification models were identified that are frequently used by researchers and outperformed the elective based classification on their respective datasets. The identified four techniques were: K-Nearest Neighbor used within collaborative filtering approach, Naïve Bayes, Decision Tree and Support Vector Machines.



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

**Decision Tree (ID3 Algorithm)**

The Decision tree (DT) is a conventional, tree shape structure which is used to determine every possible result and statistical probability, used in supervised learning as decision support tools. They contain 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 alongside. 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 as output and ten core subject and preferred subject as input. The above model had 74.5 percent accuracy.

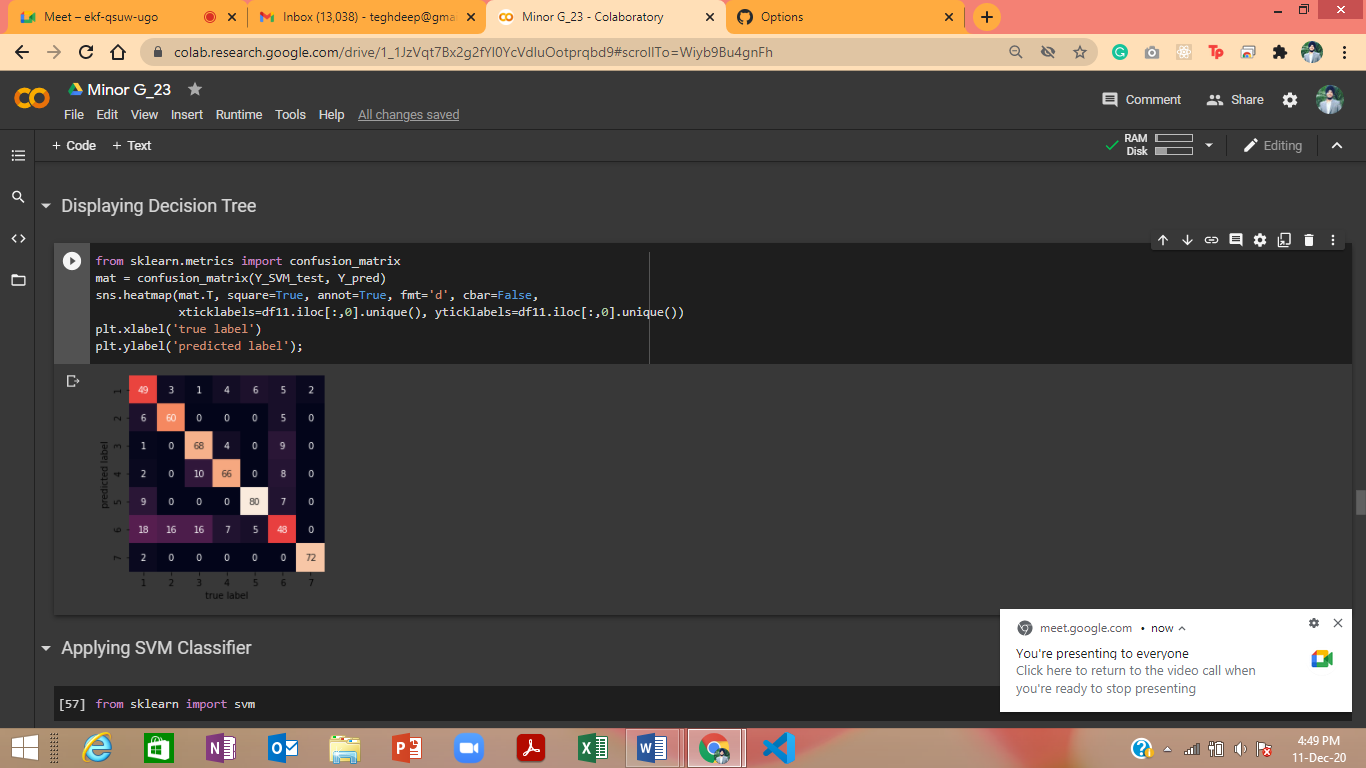


FIGURE 1: DECISION TREE

**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 of students. The above model had 77 percent accuracy.

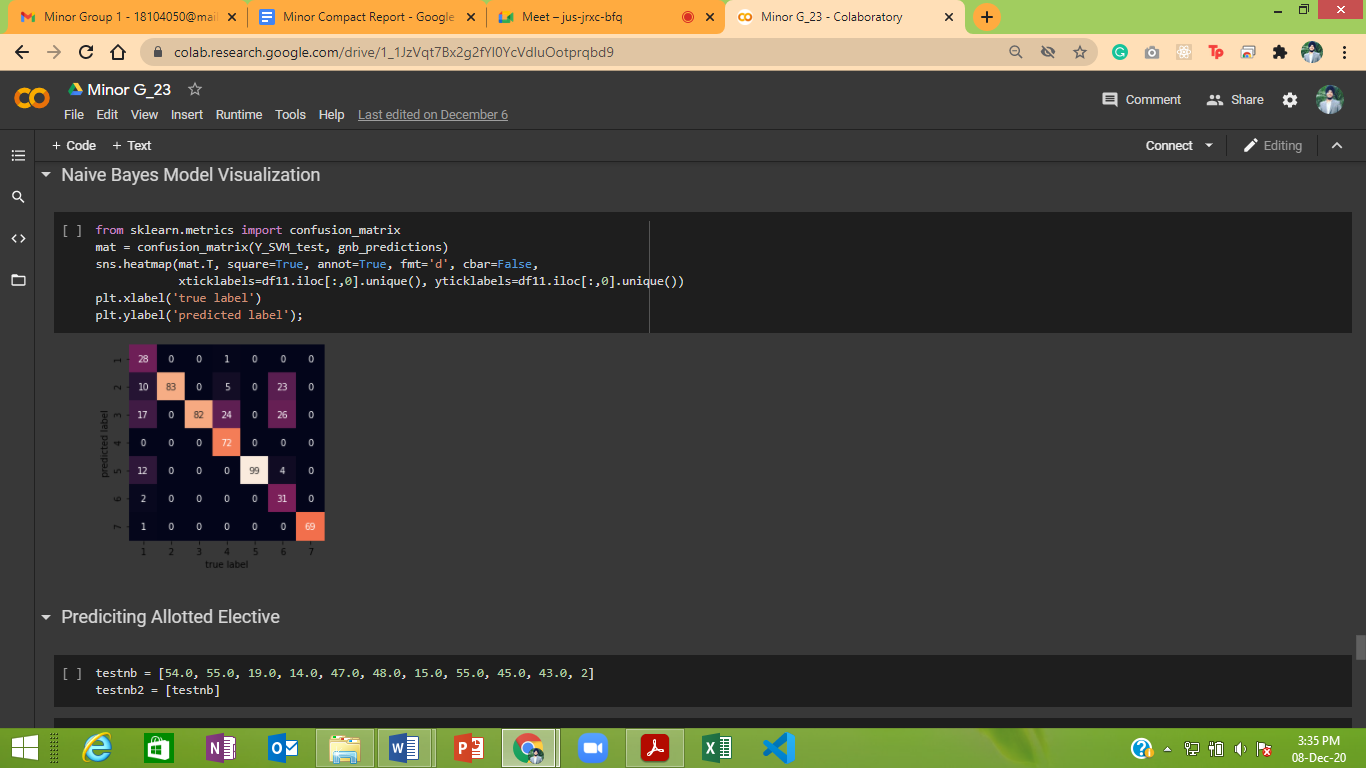


FIGURE 2: HEATMAP OF NAÏVE BAYES CONFUSION MATRIX

**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 as the output.  The above model had 72 percent accuracy

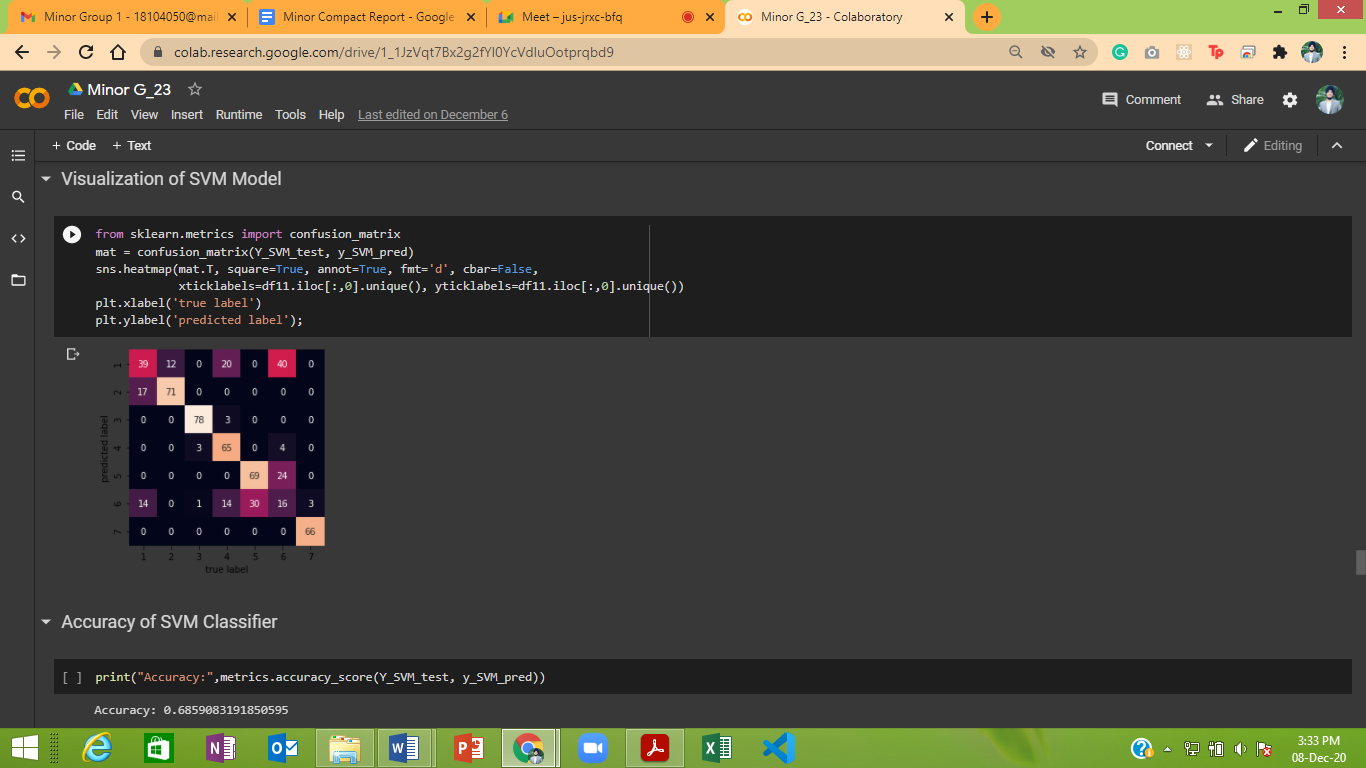


FIGURE 3: HEATMAP OF SUPPORT VECTOR MACHINES CONFUSION MATRIX

**k-Nearest Neighbor**

k-Nearest Neighbor (kNN) is lazy learning as well as a non-parametric algorithm which is used for Classification and Regression. In kNN distance measure is used to determine closeness between instances. kNN acts as a collaborative filtering technique, which stores all possible events and classifies new events based on the similarity with other similar events. Euclidean distance is used to find the separation amongst students and classified them into the similar neighbor groups. 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. It has accuracy of 22 percent.

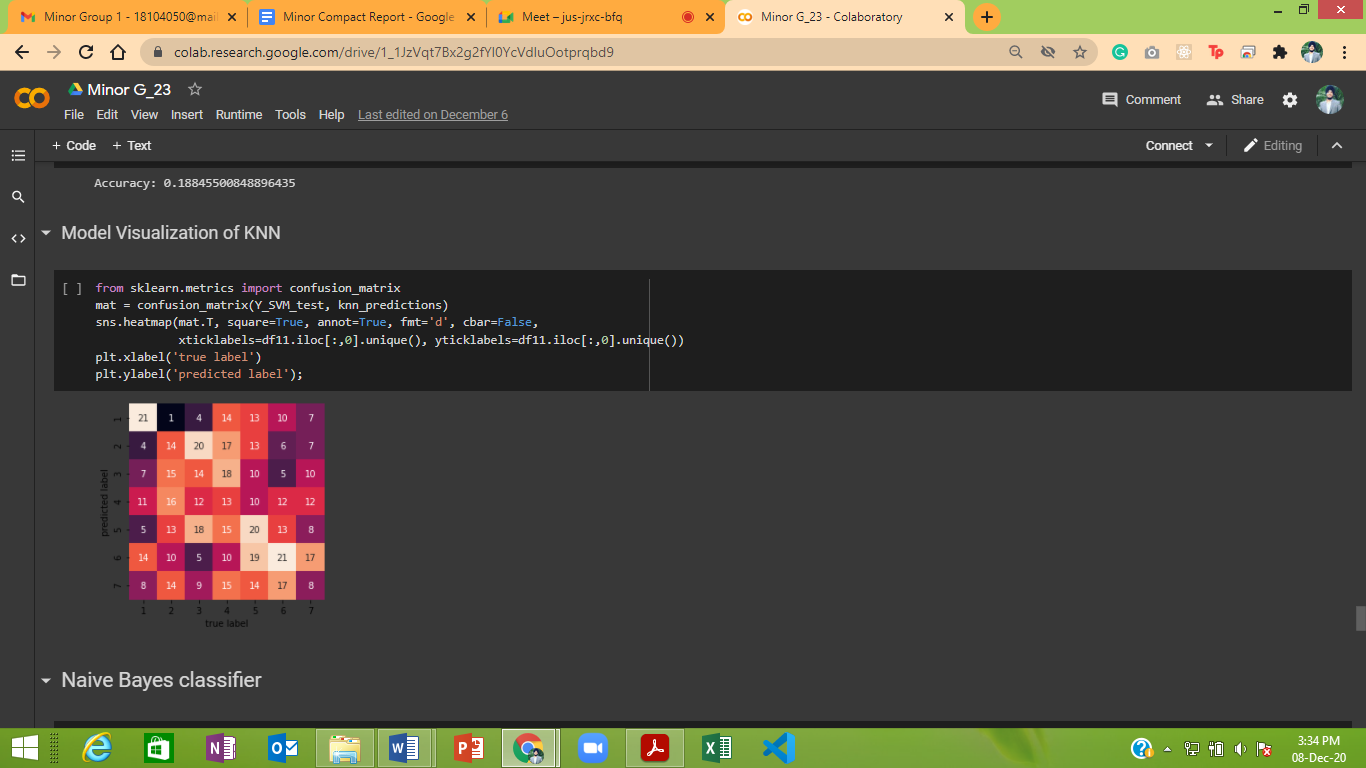


FIGURE 4: HEATMAP OF k-NEAREST NEIGHBOR CONFUSION MATRIX

**Working of Elective Recommendation Algorithm**

1. Task 1 Calculation:

User-user collaborative filtering was used to recommend users their electives on the basis of their learning pattern and scoring ability. This was done by matching the user to past data with the similar learning pattern and scoring ability. This was calculated by finding similarity between core subject marks of student and past data using Cosine Similarity. Cosine Similarity was used instead of Euclidean distance due to better results and accuracy of the model.

Data was selected if and only if it had a greater similarity score than 95% approx. Marks in each elective were calculated by taking Mean and then electives were ranked according to greater number of marks scored in the particular elective.

1. Task 2 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 Naïve Bayes was used as it was found to be the best in terms of accuracy when comparing different data mining techniques used in this study. The allotted elective was predicted by Data mining techniques and assigned appropriate weightage for final rank calculation.

1. Task 3 Calculation:

In parallel, actual percentage of students scoring a distinction (above 90%) in 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 number of distinctions scored in that particular elective.

Weighted Rank Calculation

Since we are putting every possible effort to preserve the students’ preferences along improved academic performance of the students collectively, the weighted rank concept was proposed. After the calculation of Task1, Task2 and Task3 from the above steps, final weighted rank can be calculated by assigning w1 to Task1, w2 to Task2 and w3 to Task3.

*Weighted Ranki = w1 \* Task1i + w2 \* Task2i + w3 \* Task3i*

In this study, w1 was assigned weight of 0.7, w2 was assigned weight of 0.05 and w3 weight of 0.25. However, these can further vary depending upon the institutional requirements and preferences.

Final Rank Generation

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

The subjects are then arranged in increasing order of their Final Ranks and recommended to the concerned authorities for further considerations

**RESULTS AND DISCUSSIONS**

The final results of the predictions that were obtained with the help of different models, were stored in the form of excel sheet which was further processed as per the proposed algorithm. Predictions that were 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

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.

**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)*

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.

|  |  |
| --- | --- |
| **Data Mining Models** | **Accuracy (in %)** |
| Decision Tree | 76 |
| SVM | 74 |
| Naïve Bayes | 80 |
| KNN | 22 |

Figure 1: Comparison of Accuracy in Percentage for Different Models

From this graph we can clearly see that the accuracy of Naïve Bayes model was performing better than other models that were in consideration. The Decision tree and SVM yielded somewhat comparable results and KNN had exceptionally low value in terms of accuracy.

**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)*

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 Table

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Model/Electives** | **SVV** | **CC** | **CS** | **IOT** | **ML** | **NLP** | **AR\_VR** |
| Decision Tree | 0.97 | 0.85 | 0.84 | 0.73 | 0.84 | 0.47 | 0.67 |
| SVM | 0.45 | 0.82 | 0.97 | 0.78 | 0.73 | 0.23 | 0.98 |
| Naïve Bayes | 0.95 | 0.70 | 0.54 | 1.00 | 0.85 | 0.94 | 0.96 |
| KNN | 0.35 | 0.08 | 0.22 | 0.14 | 0.09 | 0.20 | 0.21 |

Figure 2: Comparison of Precision in Percentage for Different Models

From this graph we can clearly see that Naïve Bayes was ahead of Decision Tree by fractions in terms of precision. SVM had fluctuating values, it had some highs and lows in particular electives whereas KNN had exceptionally low value in terms of precision.

**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)*

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 Table and Figure.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | **SVV** | **CC** | **CS** | **IOT** | **ML** | **NLP** | **AR\_VR** |
| Decision Tree | 0.97 | 0.78 | 0.73 | 0.83 | 0.90 | 0.56 | 0.56 |
| SVM | 0.66 | 0.91 | 0.93 | 0.66 | 0.77 | 0.15 | 1.00 |
| Naïve Bayes | 0.62 | 0.82 | 0.70 | 0.84 | 0.92 | 0.48 | 0.98 |
| KNN | 0.30 | 0.10 | 0.23 | 0.16 | 0.10 | 0.23 | 0.12 |

Figure 3: Comparison of Recall in Percentage for Different Models

From this graph we can clearly see that the accuracy of Decision tree model was performing better than other models that were in consideration. The Naïve Bayes and SVM yielded somewhat comparable results and KNN had exceptionally low value in terms of recall.

**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))*

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

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | **SVV** | **CC** | **CS** | **IOT** | **ML** | **NLP** | **AR\_VR** |
| Decision Tree | 0.97 | 0.82 | 0.78 | 0.77 | 0.87 | 0.51 | 0.61 |
| SVM | 0.54 | 0.86 | 0.95 | 0.72 | 0.75 | 0.17 | 0.99 |
| Naïve Bayes | 0.62 | 0.82 | 0.70 | 0.84 | 0.92 | 0.48 | 0.98 |
| KNN | 0.33 | 0.09 | 0.23 | 0.15 | 0.09 | 0.21 | 0.15 |

Figure 4: Comparison of F1 Score in Percentage for Different Models

**Outcomes**

After analyzing the different models used in this study on the performance evaluation parameters Naïve Bayes outperformed in terms of accuracy with an average accuracy of 78 percent across all the elective subjects. Naïve Bayes was ahead in terms of Accuracy whereas Decision Tree was ahead by fractions, in all other evaluation parameters, namely Precision, Recall and F1 Score. Except Support Vector Machines and Decision Tree that yielded somewhat comparable results, KNN was far behind in terms of accuracy, precision, recall and F1 score. An exceptional behavior of KNN which have exceptionally low values in precision, recall and F1 score parameters is also observed.

**CONCLUSION AND FUTURE SCOPE**

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, naïve bayes was found to be the best predictor classifier model for predicting the elective subject 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.

This research study has helped to fulfill the set objectives in the following ways:

* Naïve Bayes, was found to be the best predictor classifier model in the present education scenario that helped to make efficient course predictions.
* An algorithm is proposed that side-by-side considered the contextual information of individual student varying preferences.
* With the help of the predictions and individual student varying preferences, the proposed system was able to generate efficient elective course recommendations

Future Scope

The proposed elective course recommender system can be extended by:

* Consideration of cross-domain large datasets from different educational institutions for better generalized results.
* If we get a larger dataset in future, we can apply neural networks for elective recommendation.
* Other contextual attributes like gender, age, socio-economic factors etc. can also be researched upon for incorporating more efficiency within the proposed algorithm for generating more efficient elective course recommendations.

**REFERENCES**

[1] Funda Dag and Avnur Gecer,” Relations between online learning and learning styles”,

in *Proceedings of the AAAI Workshop on Educational Data Mining*, vol. 3, pp. 1-10, 2006.

[2] Cathy Li and Farah Lalani,” The COVID-19 pandemic has changed education forever. This is how”,

*Irvine, CA: University of California, School of Information and Computer Science*. 2017

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

[4] Jinjin Liang and Yong Nie,” A Hybrid Teaching Mode Based on Machine Learning Algorithm”**,**

*Research & Practice in Assessment*, vol. 9, pp. 1-9, 2014.

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

[6] D. Upendran, S. Chatterjee, S. Sindhumol and K. Bijlani, “Application of predictive analytics in intelligent course recommendation,” *Procedia Computer Science*, vol. 93, pp. 917-923, 2016.

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

[8] 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.

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

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

[11] 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,* IEEE, pp. 1-5, 2012.

[12] 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.

[13] 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.

[14] 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.

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

[16] 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.

[17] 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.

[18] 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.

[19] 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. 36-42, 2016.

**PLAGIRISM REPORT**

