Student major subject prediction model for real-application neural network

Aminul Islam [a,1,*], Jesmeen Mohd Zebaral Hoque [b,2], Md. Jakir Hossen [b,3], Halizah Basiron [a,4], Chy. Mohammed Tawsif Khan [b,5]

[a] Centre for Advanced Computing Technology, Universiti Teknikal Malaysia Melaka, Malaysia
[b] Faculty of Engineering and Technology, Multimedia University, Melaka, Malaysia
[1] aminul_islam.ai@yahoo.com; [2] jesmeen.online@gmail.com; [3] jakir.hossen@mmu.edu.my; [4] halizah@utem.edu.my; [5] tawsif.online@gmail.com
* corresponding author

ARTICLE INFO

Article history
Received October 16, 2023
Revised February 12, 2024
Accepted October 12, 2024
Available online May 31, 2025

Selected paper from The 2023 8th International Conference on Electrical, Electronics and Information Engineering (ICEEIE'23), Malang - Indonesia, September 28-29, 2023, https://iceeie.um.ac.id/. Peer-reviewed by ICEEIE'23 Scientific Committee and Editorial Team of IJAIN journal.

Keywords
Machine learning
Artificial neural network
Student's major prediction
Model selection
Admission Test Bangladesh

ABSTRACT

The university admission test is an arena for students in Bangladesh. Millions of students have passed higher secondary school every year, and only limited government medical, engineering, and public universities are available to pursue further study. It is challenging for a student to prepare all three categories simultaneously within a short period in such a competitive environment. Selecting the correct category according to the student's capability became important than following the trend. This study developed a preliminary system to predict a suitable admission test category by evaluating students' early academic performance through data collecting, data preprocessing, data modelling, model selection, and finally, integrating the trained model into the real system. Eventually, the Neural Network was selected with a maximum 97.13% prediction accuracy through a systematic process of comparing it with three other machine learning models using the RapidMiner data modeling tool. Finally, the trained Neural Network model has been implemented by the Python programming language for evaluating the possible options to focus as a major for admission test candidates in Bangladesh.

© 2025 The Author(s).
This is an open access article under the CC–BY-SA license.

1. Introduction

People can realize that education is the only way to succeed in life, and now they are more interested in educating themselves and their new generation. At present, around, on average, 78.46% of the world's population can read and write [1], but remarkable discrimination in literacy rates is observed between the states. The unemployment rate is more than 30 percent in several countries in the world, where the alarming issue is that the growth rate of literate unemployed people is increasing proportionally. Numerous literate people in South Asia are unemployed or cannot work in their field because of poor subject choice. People like to follow the ongoing trend instead of justifying their capability. As a result, they can hardly manage a job depending on what they have learned in their student life. The less literate people have a higher possibility of unemployment; as a result, they can not earn enough for themselves and their families, affecting their health and decreasing their life expectancy [2]. A well-educated person is able to earn 12% more than a poor, literate person [3]. The preliminary stage of getting a better education is admission to a good academic institution. At present, it is not easy to get admission to any reputed universities for their competitive admission test. Many students get depressed after failing

admission tests, and almost 17.7% of candidates think about suicide [4]. This study has been conducted to develop an intelligent model for admission test participants in Bangladesh to predict suitable majors for their early preparation [5].

The admission test is divided into three main categories in Bangladesh: engineering, medical, and public universities. It is nearly impossible to better prepare for all these three categories within a short time. Therefore, the students should be focused on one specific major based on their academic and personal performance. A student must pay a lot if they select the wrong category. Sometimes, they have to wait for the next year's admission test, or they lose the chance to study in government universities with reasonable tuition fees. Education in private universities is too expensive, which is not affordable for most students. Thus, it is really important to be careful of choosing a major category for admission tests. The primary goal of this study is to select and train the best-performing Machine Learning model to predict a student's major category for university admission tests in Bangladesh.

The project has been introduced by preprocessing the data. It is a lengthy and challenging process to collect students' academic performance data from all available schools in Bangladesh without proper authorization. Therefore, sample data for this project were collected from a few schools in the southwest part of Bangladesh. The collected data were processed by replacing missing and inconsistent values, removing duplicate values, feature selection, and splitting the data. The dataset contained some missing and inconsistent values in different rows. The missing values are inserted manually, replacing the inconsistent values according to the feature's characteristics. Finally, 15 features have been selected out of 22 features from the dataset, and the whole dataset is divided into two parts, with a 70:30 percent training and testing ratio. The processed datasets are applied to four selected machine learning models based on past research on students' academic performance prediction. The four models are Naive Bayes, Support Vector Machine, Artificial Neural Network, and Random Forest, where the prediction accuracies are 75.12%, 91.87%, 92.82%, and 97.13%, respectively. All four models are tested with Rapid Miner Studio. Finally, the Neural Network model was selected because of the maximum prediction accuracy. On top of that, this model is noticeably used in several studies [6]–[8] for predicting solutions. A key feature of Neural Networks is their superior training accuracy compared to linear models, particularly with large amounts of data [9]. This research aims to develop and test a Neural Network-based prediction model that can help students in Bangladesh determine the college entrance exam category that best suits their academic abilities (engineering, medical, or public universities). Thus, students can focus on preparing themselves optimally, reduce the risk of exam failure, and minimize the burden of education costs due to choosing the wrong study path.

2. Literature Review

Machine learning models are a very frequently used technique for prediction. The students' performance prediction is also a common problem that has been solved with different machine-learning techniques. Interestingly, machine learning models such as Support Vector Machine (SVM) [10]–[12], Decision Trees [13], [14], Random Forest [15], [16] Naive Bayes [17], [18] classifiers and so more are still popular among the researchers. According to [19], most student performance monitoring systems are used with a data-driven approach. They developed a double-layer algorithm with an ensemble and base predictor layers. The base predictor layer is responsible for predicting the student's regular performance related to each academic term; on the other hand, the ensemble layer tries to predict the students' posterior results by adjusting the local prediction result and the previous ensemble prediction. Side by side, a data-driven clustering method has been applied, which is based on posterior. This process is able to cluster the courses based on large, sparse, and heterogeneous student performance data. This data-driven method also applies as an alliance with different pedagogical processes for predicting student academic performance.

One closely related work was developed by [20] to predict students' acceptance or rejection of offer letters for academic admission given by a university by applying selected methods to the former students'

admission data. They obtained 66% with the selected features using k-nearest Neighbor and Decision Tree. However, they stated that using predictive models for this type of application will help the university adopt the academic admission selection technique more intelligently in the near future.

The two heavily used machine learning models, Random Forest and the Naive Bayes classifiers, can successfully predict the student's academic performance. These models are good enough for regular and small amounts of data, but the accuracy can be affected when complex data is chosen as input. The researcher only used ten different coursework numbers as input for the system, and the performance was the same for Random Forest as well as Naive Bayes Classifiers, though the random forest method is more useful in quantifying the value of different features [21].

[22] mentioned that almost thirty out of a hundred first-year students never return to continue their studies in the US. They initiated research using three machine learning techniques to predict the binary dropout variable: random forests, k-nearest neighbors, and regularized logistic regression. A regularized linear regression has been applied to predict the number of relieved semesters by each student before they drop out side by side. The K-fold cross-validation was applied to figure out the regularization potency of the system. The prediction rate of eventual student degradation yielded promising results, using information from over 32,500 students with regularized logistic regression.

The most popular machine learning model, Random Forest, can process over a million examples. Beaulac and Rosenthal (2019) [23] predicted university students' academic majors and success using RF and 1.6 million data points. RF provides variable importance metrics, aiding in understanding predictors' effects. RF is efficient, user-friendly, and outperforms linear logistic models, making it superior for diverse prediction tasks.

The projection of STEM field dropout rates for US residents due to a projected scarcity of competent professionals was covered by Aulck et al. [22]. They employed the following four machine learning methods: Logistic Regression, Random Forest, Gradient Boosted Trees, and AdaBoosted Log Reg. Except for RF, all models had similar accuracy, suggesting a possible 30% improvement in precision over the baseline ratio. Researchers promoted the pursuit of STEM degrees, looked at drivers of STEM attrition, and used a probabilistic graphical model for positive step-by-step student intention modeling in the classroom. To predict students' academic achievement, Elbadrawy developed an associative multi-regression model [24] that outperformed single linear regression by providing an accuracy gain of over 20%. Researchers used Random Forests, Personalized Multi-Linear Regression, and Factorization Machines (FM) for precise forecasting. For better performance and interpretability, they selected a combination of Random Forests and MF-based Factorization Machines, including Mean Absolute Deviation Importance (MADImp) feature selection.

Although data mining is frequently used to forecast success, more precise results may be obtained by considering well-liked machine learning techniques, including decision trees, SVM, Random Forest, and AdaBoost.R2 [25]. These models were evaluated using stratified sampling and k-fold cross-validation to account for the varying data ratios. Although there were certain patterns, there were no discernible differences in the performance of the algorithms. Predictive algorithms utilize separative and productive classification models in accordance with Daud et al. (2017) [26]. Among the five models, Classification and Regression Tree (CART), C4.5, and Support Vector Machine (SVM) are separative, whereas Bayes Network (BN) and Naive Bayes (NB) are productive. The most effective method for predicting performance is determined to be SVM.

Predicting student academic achievement depends on characteristics and prediction approaches, as [27] emphasizes. Classification techniques such as Naive Bayes, Logistic Regression, Decision Trees, K-NN, and Neural Networks are favored. Accuracy and readability must come first for a classifier to be effective. Researchers discovered that Naive Bayes excels with more accuracy among the selected algorithms.

3. Method

The Complete process of developing the system is divided into two sections, i.e., data preparation, Modeling, and Evaluation. Data preparation and understanding consists of data gathering, discovering and assessing data, transforming and enriching data, cleansing and validating data, and data storing. Secondly, the identification of suitable models for the data set is processed by evaluating the optimistic one.

3.1. Data Preparation

3.1.1. Data Collection

The data of this project has been collected from some specific schools and colleges in the southwest part of Bangladesh through a shadow educational institution, Spark Academic and Admission Care. The data was collected from students who have completed higher secondary school within 5 years and were admitted into any public university, engineering university, or government medical college. The data set is named SPARK data and listed by a mixed method, which contains some numerical as well as some words and names. This is a secondary data type, but it has not been used before. These data are collected from the administration records. First of all, students are directly asked to fill out a structured form to confirm their basic information and secondary school results during their registration process into the academy. However, the institution agreed to use only the academic information to predict student academic performance. Student's personal information was not recorded for this study. The dataset contains students' secondary and higher secondary academic results of each course according to the institution names, their gender and roll number, and so on. This data set has 22 features along with almost 700 data, where 500 data have been used to train the system, and the rest are for testing the system's performance.

3.1.2. Data Preprocessing

Data preprocessing is an important stage for training and Machine Learning (ML) models. Developing the data quality by extracting meaningful insights from the dataset is helpful. This is the primary phase where all the data has been prepared for the training session. First of all, the missing data has been removed or exchanged with the new value to ensure the dataset is filled up properly. Next, irrelevant attributes in datasets are moved; for instance, in this research, students' parents' names, contact numbers or email IDs, favorite color or food, whether they have any pets, their height, screen color, etc., are not considered. So, the appropriate and relevant attribute can increase the system performance by reducing run-time. Labelling is another prepossessing activity where all the feature's name has been changed into meaningful value if they have not been done before.

3.1.3. Feature selection

The data collection is a very difficult and hard-working process. Thus, the researchers try to collect all kinds of data whenever they start a survey. All those data might not be necessary for that particular project, but they reserve it as secondary data. Here, the data set of Spark Academic and Admission Care consists of 22 features and 700 examples. All these 22 features are not important in predicting students' academic performance. Therefore, the feature has to be evaluated to get the best result. Finally, 15 important features for the training process, according to the literature review which, are S_Math, S_Phys, S_ICT, S_Eng, S_Chem, S_Bio, S_GPA, H_Math, H_Phys, l H_ICT, H_Eng, H_Chem, H_Bio, H_GPA, Adm_Inst was selected as show in Fig. 1.

3.1.4. Splitting the Dataset

The data is split randomly into training and testing parts to evaluate the selected ML models. The most popular splitting ratio for the datasets is 70:30 or 80:20, where either 70% or 80% of the data is allocated to the training set, and the remaining 30% or 20% of the data is reserved for testing purposes. The 70:30 splitting ratio has been used for this dataset to train and test the model, as it has a better performance ratio than other [28], as shown in Fig. 2.

1	S_Math	S_Phys	S_ICT	S_Eng	S_Chem	S_Bio	S_GPA	H_Math	H_Phys	H_ICT	H_Eng	H_Chem	H_Bio	H_GPA	Adm_Inst
2	81	83	81	84	90	87	5	83	84	82	84	91	92	5	M
3	90	76	92	86	84	80	5	86	86	85	83	85	75	5	E
4	91	75	82	83	82	84	5	85	83	84	83	85	86	5	E
5	87	85	86	86	81	80	5	87	87	86	85	85	86	5	E
6	82	86	88	88	87	92	5	77	83	81	87	86	95	5	M
7	72	75	80	86	90	84	4	81	77	84	87	84	90	5	P
8	74	87	63	87	75	91	4	80	81	82	84	90	93	5	M
9	80	82	81	86	87	92	5	62	81	71	87	85	84	4	M
10	83	85	82	77	83	83	5	86	86	76	84	87	83	5	E
11	86	86	90	75	76	83	4	90	87	86	77	81	85	5	P

Fig. 1. A snippet of the final dataset with the selected features

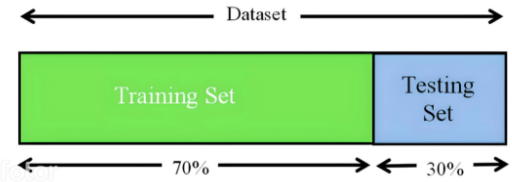

Fig. 2. Splitting ratio of the whole dataset to train and test the model

3.2. System Modeling

According to the dataset, different ML models perform differently. This section presents the implementation of various models using data modeling techniques with RapidMiner Studio. The main goal is to identify the best-suited model for the dataset. Here, the four most frequently practiced ML models for students' academic performance prediction are selected through past research.

3.2.1. Artificial Neural Network (ANN)

The Neural Network (NN) algorithm mimics the structure and operation of organic neurons. It utilizes an activation function or threshold function to make decisions as it examines nominal inputs to determine the closest match. The activation function of each node in an ANN includes Binary Step, Linear, Sigmoid, ReLU, Swish, and SoftMax, among others. Input, output, and hidden layers are present in ANNs [29]. There may be one or more hidden layers. Neurons send signals to other neurons or, in the case of back-propagation, to the preceding node for weight modification. The input (x) and weight (w) of a neuron are given by the equation (1).

$$z = x.w + b \tag{1}$$

where, the b is the bias or offset. It helps to adjust the value based on the output, and the weight indicates the connection strength between neurons, as well as the influence the given input will have on the output. If a neuron takes multiple inputs, then the equation is defined as (2).

$$z = \sum_{i=0}^{n} x_i w_i + b \tag{2}$$

here, n is the number of input nodes.

3.2.2. Naive Bayes (NB)

The Naive Bayes classification is a Bayes theorem-based algorithm that can build an excellent model even with a small data set. This is a high-bias and low-variance classifier. The NB is developed using a group of algorithms, rather than a single algorithm, but all share a common principle. Naive Bayes is one of the easiest models to generate, but it can fit almost all sizes of data sets. Bayes' theorem provides a way of calculating posterior probability as shown in (3).

$$P(m|n) = \frac{P(n|m)\,P(m)}{P(n)} \quad (3)$$

here, $P(m|n)$ is the probability of an event m occurring relative to the event n where the event n is already given. On the other hand, the $P(n|m)$ is the reverse of $P(m|n)$, where the probability is measured based on the event m. The $P(m)$ and $P(n)$ are the chances of the particular event based on given data.

3.2.3. Random Forest (RF)

RF mixes different random trees, where each tree has leaves that indicate potential outcomes and nodes that represent input factors. In RF, trees are arranged from the root to the leaf. In contrast to an empty field, when hitting even one tree is uncommon, striking at least one tree in a forest is far more likely. Entropy and Gini impurity are two impurity criteria that RF uses for classification jobs. The Gini index is calculated as (4), and entropy is defined as (5) for a dataset with C classes.

$$gini(C) = \sum_{i=1}^{n_D} f_i\,(1 - f_i) \quad (4)$$

$$ntropy(C) = \sum_{i=1}^{n_D} -f_i\,log(f_i) \quad (5)$$

Here, the nD is the number of unique labels of the dataset, and fi is the frequency of label i at node. If the dataset split into more than one class with R1, R2, ..., Rn ratio, then the Gini index would be like (6).

$$gini_{multi}(C) = \frac{R_1}{R} gini(C_1) + \frac{R_2}{R} gini(C_2) + \ldots + \frac{R_n}{R} gini(C_n) \quad (6)$$

3.2.4. Support Vector Machine (SVM)

This flexible approach performs well on a variety of learning tasks and is used for both classification and regression. Based on distinctive traits, it projects data into an n-dimensional space and identifies a hyperplane to divide classes. Due to its versatility in handling both linear and non-linear separable problems as well as high-dimensional data, linear SVMs are well-liked in the fields of pattern recognition and classification. They divide the data using the decision boundary, which is represented by (7).

$$f(x) : w^T x + b = 0 \quad (7)$$

Here, the w represents the decision boundary normal vector, and the b perpendicular distance from the origin to the decision boundary line.

Data modeling is followed by model completion. In-depth data understanding and storage are the first steps in data modeling, which is then followed by testing several ML models and choosing the best one. Different machine learning models may be applied and compared with the help of data modeling tools like RapidMiner. The finished trained model is then saved and used in practical applications.

Three categories—multiclass classification, binary classification, and regression—were used to evaluate the prediction algorithm. Although NB and SVM don't enable regression, they perform well in binary and multiclass classification. With the dataset divided into engineering, medicine, and public universities, RapidMiner may be used to apply ANN, NB, RF, and SVM for comparison. The Naive Bayes implementation in RapidMiner is shown in Fig. 3. NB demonstrates high accuracy, even with small datasets, making it a high-bias and low-variance classifier.

The implementation process does not require any hard coding. Firstly, SPARK data is retrieved and uploaded into the repository. Next, by using the Set Role option, the goal/target of the model to train is set. The next step is to divide the data set into training and testing sets by using a split data operator. The next stage connects the NB modeling operator with the training data set. The applied model operator helps to connect the generated model with the testing data set. Finally, the performance operator helps to calculate the accuracy of the model.

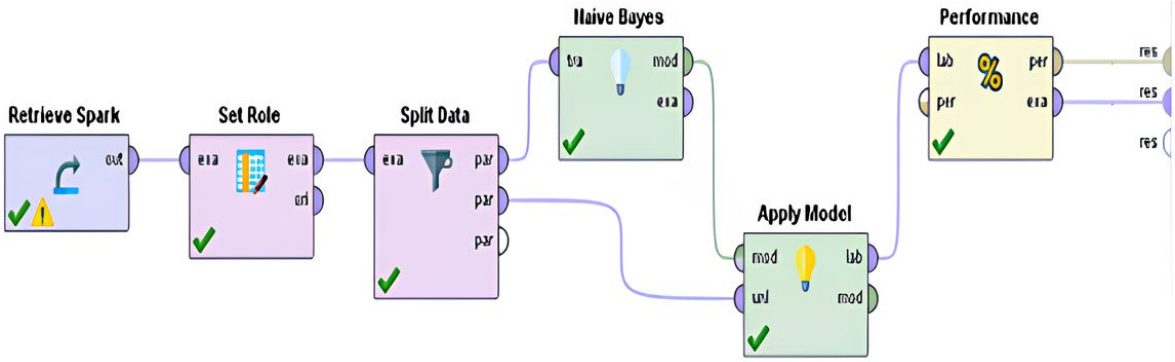

Fig. 3. The process of implementing Naive Bayes in RapidMiner

For RF, the NB operator has to be replaced by an RF modeling operator. RapidMiner provides a parameter called "Number of Trees" to specify the size. Fig. 4 discloses the trained sample using RF. Each attribute is differentiated by a splitting rule, represented by a single node in the tree. Next, to integrate SVM in this flow, a new operator is required to add for Nominal to Numerical conversion before setting the role. The SVM algorithm does not support polynomial values. The data type of the target attribute is polynomial because it contains the values M, E, and P. Therefore, the Nominal to Numerical operator was used here to convert the polynomial value into numerical by replacing the M, E, and P with 0, 1, and 2, respectively.

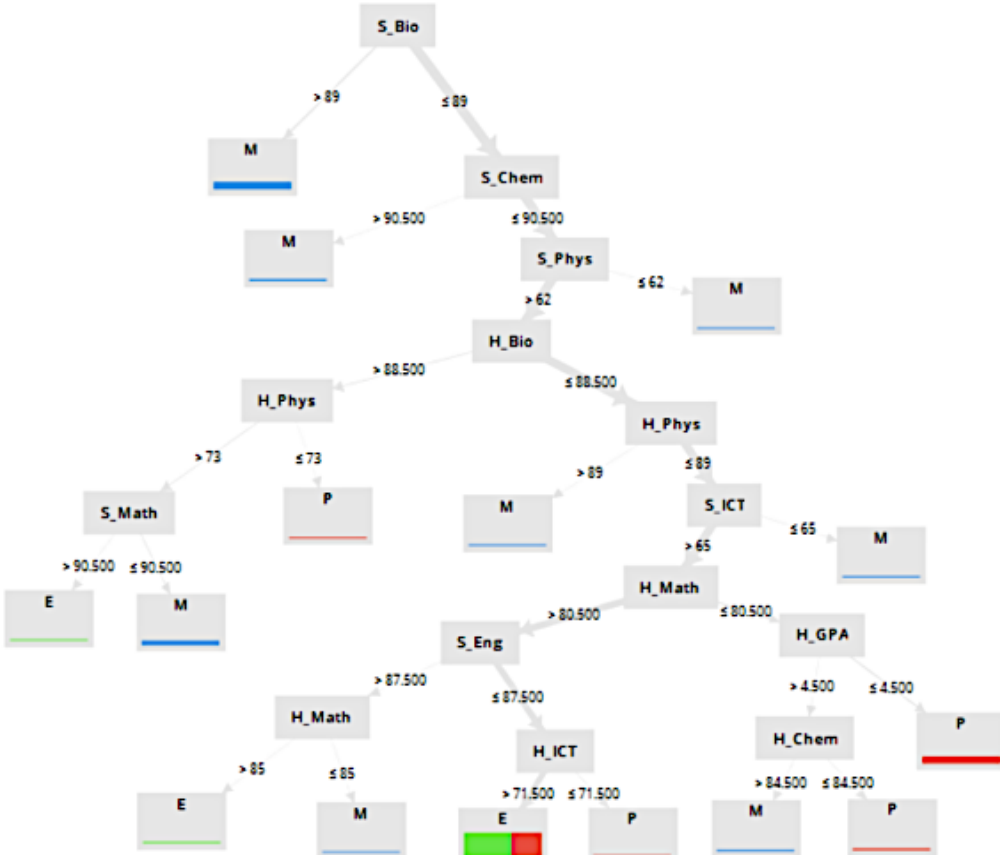

Fig. 4. A sample result of Random Forest

Replacing the NB operator with ANN was tested without needing a Nominal to Numerical operator, as only the target attribute was polynomial. Default parameters were used, with a training cycle of 200 iterations, a learning rate of 0.01, and momentum set to 0.9 to optimize and prevent local maxima. With these settings, the system was ready for execution, producing results akin to those in Fig. 5.

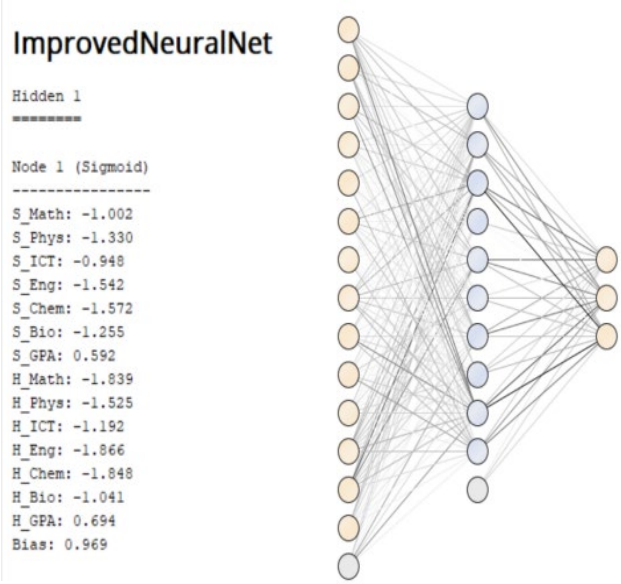

Fig. 5. A sample output of Neural Network

4. Results and Discussion

The four selected machine learning models have been tested with the selected data set. However, only one model can be selected at a time for implementation. Thus, the best-performed model for the dataset can be chosen by comparing all four results.

First, the NB comes with 75.12% accuracy, as shown in Table I. It can successfully predict 48 Medical students, 53 Engineering students, and 56 public university students out of 56, 71, and 82 students, respectively. Therefore, Naive Bayes can successfully predict a total of 157 students out of 209 students from three different categories. This is a considerable prediction, but we cannot consider it good accuracy.

Table 1. Performance of Naive Bayes

	True M	True E	True P	Class Precision
Predict M	48	0	6	88.89%
Predict E	1	53	20	71.62%
Predict P	7	18	56	69.14%
Class recall	85.71%	74.65%	68.29%	

Next, Table 2 shows the 91.87% noticeable accuracy of SVM. As it discloses, the model can successfully predict 51 Medical students, 68 Engineering students, and 73 public university students out of 56, 71, and 82 students, respectively. Therefore, the support vector machine can successfully predict a total of 192 students out of 209 students from three different categories. This is a better prediction rate than the Naive Bayes model.

Table 2. Performance of Support Vector Machine

	True M	True E	True P	Class Precision
Predict M	51	0	0	100%
Predict E	4	68	9	83.95%
Predict P	1	3	73	94.81%
Class recall	91.07%	95.77%	89.02%	

The RF is the third predictive mode, which has 92.82% accuracy. Table 3 refers to the performance of the RF model against the data set. This model can successfully predict 53 Medical students, 68 Engineering students, and 73 public university students out of 56, 71, and 82 students, respectively. Therefore, the SVM can successfully predict a total of 192 students out of 209 students from three different categories, as it is noticed that the prediction rate is almost the same for RF and SVM.

Table 3. Performance of Random Forest

	True M	True E	True P	precision
Predict M	53	0	0	100%
Predict E	0	68	9	88.31%
Predict P	3	3	73	92.41%
Class recall	94.64%	95.77%	89.02%	

The final selected ML model is the ANN. It successfully predicted a 97.13% accurate solution for the dataset, shown in Table 4. This model can predict the performance of 56 Medical students, 69 Engineering students, and 78 public university students out of 56, 71, and 82 students, respectively. Therefore, the ANN misses only 6 out of 209 examples from three categories. This is a good prediction, with the best prediction rate among all four models.

Table 4. Performance of Neural Network

	True M	True M	True P	Precision
Predict M	56	0	0	100%
Predict E	0	69	4	94.52%
Predict P	0	2	78	97.50%
Class recall	100%	97.18%	95.12%	

The four ML models were evaluated using the ROC graph. Fig. 6 shows the ROC comparison for all four ML models.

Here, the NB, SVM, ANN, and RF are respectively marked by blue, green, yellow, and red color lines. However, it is observed that the ANN shows better performance than all other models, where the NB is the poorest one. The SVM and the RF show almost identical performance by overlapping the lines.

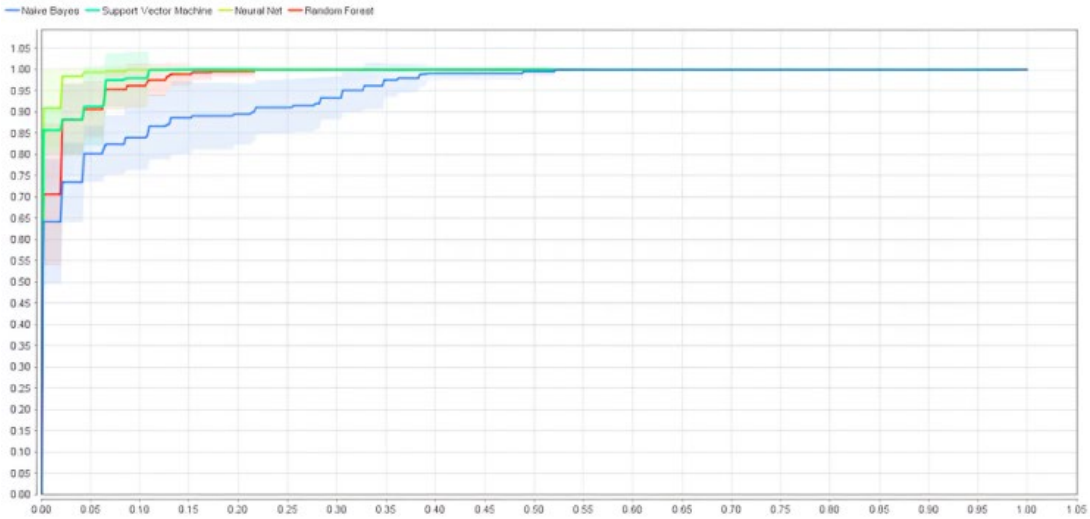

Fig. 6. The ROCs Comparison for selected machine learning models

In this case, the four selected machine learning models are frequently used to solve different prediction problems, even though all models show more than 70% prediction accuracy in this study.

However, applying some of the four models at a time is only possible. Therefore, the Neural Network model has been selected for implementation as it depicts better prediction accuracy compared with the other three machine learning models.

5. Conclusion

In this paper, a neural network model has been systematically selected and trained with the predefined training data set to predict a suitable university admission test category for candidates in Bangladesh, which can help reduce the tuition cost and study gap among students. The Neural Network was selected by data modeling process through comparison with Naive Bayes, Random Forest, and Support Vector Machine, which performed the highest 97.13% accuracy. The trained model was evaluated with the test data set and successfully predicted the 98% results of the admission test category correctly. In this case, this model is trained with a few numbers of data and small diversity, but it is performing well by meeting the objectives. However, this system is applicable in a specific area only. In the future, this project can be extended nationwide by training it with a big dataset. The dataset has to be collected from different regions of the country to maintain data diversity. Every system needs to be maintained throughout its lifelong service, either to develop and upgrade the system or to fix the error. This system also has the same life cycle. In this case, the main improvement would be upgrading the training dataset. The data has been collected from several academic institutions in the southwest part of Bangladesh under the same education board. Therefore, the trained model is only capable of performing a good accuracy only in that specific area. However, this problem can easily be overcome by gathering enough training data from all over the country to ensure data equitability. The data must be collected from all eleven education boards in Bangladesh and all schools under those boards. Some other features can also be added to the dataset, like the name of the education boards and schools, school ranking, student performance in schools, gender, the location of the institution, whether it is situated in a town or village, etc. Nevertheless, this is a complicated data collection process, but possible. Thus, this model can be applied to predict the university admission test category for candidates in some specific areas of Bangladesh; however, it is possible to initialize the project all over the country by improving the training data.

Declarations

Author contribution. All authors contributed equally to the main contributor to this paper. All authors read and approved the final paper.
Funding statement. None of the authors have received any funding or grants from any institution or funding body for the research.
Conflict of interest. The authors declare no conflict of interest.
Additional information. No additional information is available for this paper.

References

[1] "Literacy Rate by Country 2025," *World Population Review*. [Online]. Available at: https://worldpopulationreview.com/country-rankings/literacy-rate-by-country.

[2] "Literacy and life expectancy," *National Literacy Trust*, 2024. [Online]. Available at: https://literacytrust.org.uk/research-services/research-reports/literacy-and-life-expectancy/.

[3] J. Bynner, "Improving adult basic skills: benefits to the individual and to society," DfEE Publications, pp. 1-88, 2001. [Online]. Available: https://www.voced.edu.au/content/ngv%3A35806.

[4] M. A. Mamun, J. M. Misti, I. Hosen, and F. Mamun, "Suicidal behaviors and university entrance test-related factors: A Bangladeshi exploratory study," *Perspect. Psychiatr. Care*, vol. 58, no. 1, pp. 278–287, Jan. 2022, doi: 10.1111/ppc.12783.

[5] Neha, A. Sharma, and R. Kaur, "Recommendation System For Course Selection In Higher Education: A Critical Survey," *Section A-Research paper*, pp. 6874 - 6886, 2023. [Online]. Available at: https://www.coursehero.com/file/232321335/505903c3b843c2f5b2f165f2bd49dee1pdf/.

[6] C. F. Rodríguez-Hernández, M. Musso, E. Kyndt, and E. Cascallar, "Artificial neural networks in academic performance prediction: Systematic implementation and predictor evaluation," *Comput. Educ. Artif. Intell.*, vol. 2, p. 100018, Jan. 2021, doi: 10.1016/j.caeai.2021.100018.

[7] C. Huang, J. Zhou, J. Chen, J. Yang, K. Clawson, and Y. Peng, "A feature weighted support vector machine and artificial neural network algorithm for academic course performance prediction," *Neural Comput. Appl.*, vol. 35, no. 16, pp. 11517–11529, Jun. 2023, doi: 10.1007/s00521-021-05962-3.

[8] A. Rivas, A. González-Briones, G. Hernández, J. Prieto, and P. Chamoso, "Artificial neural network analysis of the academic performance of students in virtual learning environments," *Neurocomputing*, vol. 423, pp. 713–720, Jan. 2021, doi: 10.1016/j.neucom.2020.02.125.

[9] R. Dastres and M. Soori, "Artificial Neural Network Systems," *Int. J. Imaging Robot.*, vol. 2021, no. 2, pp. 13–25, 2021. [Online]. Available at: www.ceserp.com/cp-jour.

[10] M. T. Tally and H. Amintoosi, "A hybrid method of genetic algorithm and support vector machine for intrusion detection," *Int. J. Electr. Comput. Eng.*, vol. 11, no. 1, p. 900, Feb. 2021, doi: 10.11591/ijece.v11i1.pp900-908.

[11] M. Zulqarnain, R. Ghazali, Y. M. Mohmad Hassim, and M. Rehan, "Text classification based on gated recurrent unit combines with support vector machine," *Int. J. Electr. Comput. Eng.*, vol. 10, no. 4, p. 3734, Aug. 2020, doi: 10.11591/ijece.v10i4.pp3734-3742.

[12] M. Farhadi and N. Mollayi, "Application of the Least Square Support Vector Machine for point-to-point forecasting of the PV Power," *Int. J. Electr. Comput. Eng.*, vol. 9, no. 4, p. 2205, Aug. 2019, doi: 10.11591/ijece.v9i4.pp2205-2211.

[13] S. Khan and K. V. B, "Comparing machine learning and ensemble learning in the field of football," *Int. J. Electr. Comput. Eng.*, vol. 9, no. 5, p. 4321, Oct. 2019, doi: 10.11591/ijece.v9i5.pp4321-4325.

[14] R. R N, V. R., and A. M. R., "Autonomous Traffic Signal Control using Decision Tree," *Int. J. Electr. Comput. Eng.*, vol. 8, no. 3, p. 1522, Jun. 2018, doi: 10.11591/ijece.v8i3.pp1522-1529.

[15] M. A. Islam Arif, S. I. Sany, F. Sharmin, M. S. Rahman, and M. T. Habib, "Prediction of addiction to drugs and alcohol using machine learning: A case study on Bangladeshi population," *Int. J. Electr. Comput. Eng.*, vol. 11, no. 5, p. 4471, Oct. 2021, doi: 10.11591/ijece.v11i5.pp4471-4480.

[16] R. S and S. Kumar J, "Performance evaluation of random forest with feature selection methods in prediction of diabetes," *Int. J. Electr. Comput. Eng.*, vol. 10, no. 1, p. 353, Feb. 2020, doi: 10.11591/ijece.v10i1.pp353-359.

[17] A. Triayudi, S. Sumiati, S. Dwiyatno, D. Karyaningsih, and S. Susilawati, "Measure the effectiveness of information systems with the naïve bayes classifier method," *IAES Int. J. Artif. Intell.*, vol. 10, no. 2, p. 414, Jun. 2021, doi: 10.11591/ijai.v10.i2.pp414-420.

[18] T. Winarti, H. Indriyawati, V. Vydia, and F. W. Christanto, "Performance comparison between naive bayes and k- nearest neighbor algorithm for the classification of Indonesian language articles," *IAES Int. J. Artif. Intell.*, vol. 10, no. 2, p. 452, Jun. 2021, doi: 10.11591/ijai.v10.i2.pp452-457.

[19] J. Xu, K. H. Moon, and M. van der Schaar, "A Machine Learning Approach for Tracking and Predicting Student Performance in Degree Programs," *IEEE J. Sel. Top. Signal Process.*, vol. 11, no. 5, pp. 742–753, Aug. 2017, doi: 10.1109/JSTSP.2017.2692560.

[20] M. Y. Iqbal Basheer, S. Mutalib, N. Hamimah Abdul Hamid, S. Abdul-Rahman, and A. M. Ab Malik, "Predictive analytics of university student intake using supervised methods," *IAES Int. J. Artif. Intell.*, vol. 8, no. 4, p. 367, Dec. 2019, doi: 10.11591/ijai.v8.i4.pp367-374.

[21] Q. Wu, "Predicting Academic Performance Via Machine Learning," Texas A&M University, pp. 1-17, 2017. [Online]. Available at: https://oaktrust.library.tamu.edu/items/0a577649-066e-4e10-ba1f-22cb4299e273.

[22] L. Aulck, R. Aras, L. Li, C. L'Heureux, P. Lu, and J. West, "Stem-ming the Tide: Predicting STEM attrition using student transcript data," *arXiv*, no. August, pp. 1–10, 2017, [Online]. Available at: https://arxiv.org/abs/1708.09344.

[23] C. Beaulac and J. S. Rosenthal, "Predicting University Students' Academic Success and Major Using Random Forests," *Res. High. Educ.*, vol. 60, no. 7, pp. 1048–1064, Nov. 2019, doi: 10.1007/s11162-019-09546-y.

[24] A. Elbadrawy, R. S. Studham, and G. Karypis, "Collaborative multi-regression models for predicting students' performance in course activities," in *Proceedings of the Fifth International Conference on Learning Analytics And Knowledge*, Mar. 2015, vol. 16-20-Marc, pp. 103–107, doi: 10.1145/2723576.2723590.

[25] P. Strecht, L. Cruz, C. Soares, J. Mendes-Moreira, and R. Abreu, "A Comparative Study of Classification and Regression Algorithms for Modelling Students' Academic Performance.," in *International Educational Data Mining Society*, Jun. 2015, pp. 392–395. [Online]. Available at: https://eric.ed.gov/?id=ED560769.

[26] A. Daud, N. R. Aljohani, R. A. Abbasi, M. D. Lytras, F. Abbas, and J. S. Alowibdi, "Predicting Student Performance using Advanced Learning Analytics," in *Proceedings of the 26th International Conference on World Wide Web Companion - WWW'17 Companion*, 2017, pp. 415–421, doi: 10.1145/3041021.3054164.

[27] W. F. W. Yaacob, S. A. M. Nasir, W. F. W. Yaacob, and N. M. Sobri, "Supervised data mining approach for predicting student performance," *Indones. J. Electr. Eng. Comput. Sci.*, vol. 16, no. 3, p. 1584, Dec. 2019, doi: 10.11591/ijeecs.v16.i3.pp1584-1592.

[28] Q. H. Nguyen *et al.*, "Influence of Data Splitting on Performance of Machine Learning Models in Prediction of Shear Strength of Soil," *Math. Probl. Eng.*, vol. 2021, no. 1, p. 4832864, Jan. 2021, doi: 10.1155/2021/4832864.

[29] M. R. Siregar, A. Rifai, and M. Mariana, "Overcoming the Global Economic Crisis in The Perspective of Islamic Finance," *Al-Kharaj J. Islam. Econ. Bus.*, vol. 5, no. 3, pp. 406–417, 2023, doi: 10.24256/kharaj.v5i3.4146.