

Predicting Students' Performance of the Private Universities of Bangladesh using Machine Learning Approaches

Md. Sabab Zulfiker¹, Nasrin Kabir², Al Amin Biswas³, Partha Chakraborty⁴, and Md. Mahfujur Rahman⁵
Department of Computer Science and Engineering, Daffodil International University, Dhaka, Bangladesh^{1,2,3,5}
Department of Computer Science and Engineering, Comilla University, Cumilla, Bangladesh⁴

Abstract—Every year thousands of students get admitted into different universities in Bangladesh. Among them, a large number of students complete their graduation with low scoring results which affect their careers. By predicting their grades before the final examination, they can take essential measures to ameliorate their grades. This article has proposed different machine learning approaches for predicting the grade of a student in a course, in the context of the private universities of Bangladesh. Using different features that affect the result of a student, seven different classifiers have been trained, namely: Support Vector Machine (SVM), K-Nearest Neighbor (KNN), Logistic Regression, Decision Tree, AdaBoost, Multilayer Perceptron (MLP), and Extra Tree Classifier for classifying the students' final grades into four quality classes: Excellent, Good, Poor, and Fail. Afterwards, the outputs of the base classifiers have been aggregated using the weighted voting approach to attain better results. And here this study has achieved an accuracy of 81.73%, where the weighted voting classifier outperforms the base classifiers.

Keywords—Prediction; machine learning; weighted voting approach; private universities of Bangladesh

I. INTRODUCTION

Nowadays various statistical and machine learning algorithms are applied in different fields, such as marketing, health and medical issues, weather forecasting, socioeconomic behavior analysis, etc. It has emerged to educational data also. From the perspective of Bangladesh, the number of the private universities is much more than the number of the public universities. Currently in Bangladesh, there are one hundred and five private universities [1]. As a result, the number of students in the private universities is much higher than the public universities. Since we cannot imagine the development of the higher education except the development of the quality of the private universities, it is necessary to focus on the students of the private universities.

The databases of the different universities store a large volume of data. This data include the data of the students, teachers, and employees of the universities. By analyzing this data, different patterns can be derived which will be helpful to make decisions. Using diverse machine learning and data mining techniques on these data, many kinds of knowledge can be discovered and this knowledge can be used to predict the enrolment status of the students in a course, to detect illegal activities in the online examination and, to identify unusual marks in the result sheet, etc. [2]. Different statistical analysis and machine learning algorithms can be applied on the data of the students of the universities for predicting the grades of

different courses that they have taken in their undergraduate level. There is a massive growth in the number of students who are getting admitted in different public and private universities of Bangladesh. A vast portion of these students can not gather proper skills and knowledge in their four years' tenure of the university life. Not only practical knowledge, a huge number of students come out of universities with low scoring results. As they lack both theoretical knowledge and practical knowledge, it becomes very hard for them to place themselves in job markets. If the students can predict their grades or results before their final examinations, they can take necessary actions to improve their results. Then the teachers can also identify which students are at risk and so they can guide the weak students properly and help them to recover [3]. For predicting the performance of the students predictive modeling can be used. Several methods can be used for building a predictive model like: classification, regression, categorization. Among these methods, classification has the most popularity [4].

After the accomplishment of this research, it will help to find out the different approaches to predict the students' final grades as well as determine the best approach for performing the prediction.

The main objectives of this research are: to predict the final grades of the students in a course using different machine learning algorithms, to forecast whether a student is at risk of failure in the final examination or not, and to compare the results of different machine learning algorithms for identifying which algorithm gives the best performance.

The residue of the paper is structured as follows: Section II describes the related works, Section III exhibits the entire methodology. The results are discussed in Section IV, the conclusion is discussed in Section V. Finally, Section VI represents the future works.

II. RELATED WORKS

Yadav and Pal [2] performed a study on predicting the results of the students of 1st year of Engineering. They collected data from the enrolment form which were filled by the students during their admission in VBS Purvanchal University, Jaunpur. With this dataset, they built models using different variations of the Decision Tree algorithm for classifying the students' performances in the year final examination of the first year. They showed that C4.5 obtained the highest accuracy of 67.78%.

Kabra and Bichkar [3] collected data from the entry form, filled by the students in an engineering college during the time of admission. Using J48 algorithm, they predicted the final grades of the first year students'. When they classified the results into three categories they gained an accuracy of 60.46% and in the case of classifying the results into two categories they gained an accuracy of 69.94%.

Kapur et al. [5] used various classification algorithms to classify the performance of the students into three categories: high, medium, and low. Their dataset included 480 entries with 16 attributes. Among these classifiers Random Forest showed the highest accuracy of 76.67%.

Liu and Zhang [6] gathered 210 records of the students. The dataset contained the marks of some major subjects and with this dataset, they trained C4.5 classifier for predicting whether a student would pass or fail.

Sweeney et al. [7] proposed a system for predicting the grades of the students for the next enrollment term. They applied two classes of methods: Simple Baselines and Matrix Factorization (MF) based methods. The lowest prediction error was achieved by the Factorization Machine (FM) Model of Matrix Factorization based methods.

Yadav et al. [8] performed a comparative study among the CART, C4.5, and ID3 algorithms for predicting the end semester marks. The dataset contained a variety of attributes like: marks achieved in the last semester, grades obtained in the class test, attendance marks, lab work performances, etc. They used WEKA explorer as the data mining tool.

Minaei-Bidgoli et al. [9] performed their study on LON-CAPA, which is an online education system. Firstly, for the classification purpose they used diverse base classifiers like: Parzen window, 1- Nearest Neighbor (1NN), K-Nearest Neighbor (KNN), Quadratic Bayesian Classifier, Multilayer Perceptron (MLP), and Decision Tree. For improving the accuracy, they also made use of a combination of the classifiers. Finally, for optimizing the accuracy of the combination of the classifiers they used Genetic Algorithm (GA). They found that Genetic Algorithm increased the accuracy by 10-12%.

Z. Iqbal et al. [10] found that the CGPA of a student in the degree program is high, if his university entry test score and HSSC (Higher Secondary School Certificate) score is high. They compared the performance of Restricted Boltzmann Machine (RBM), Matrix Factorization (MF), Collaborative Filtering (CF) and showed that RBM exhibited the best performance.

A comparative study between four distinct models: Step-wise Polynomial Regression, Linear Decision Rule, Linear Multiple Regression, and a simple Artificial Neural Network were proposed by Gorr et al. for predicting students' GPA [11].

Meier et al. [12] stated that the timely prediction of the final grade is also significant. So they proposed an algorithm that could not only predict the final grades but also performed timely prediction using previous performances of the students.

Jishan et al. [13] showed that preprocessing the data with the combination of Synthetic Minority Over-Sampling and Optimal Equal Width Binning significantly improves the accuracy of predicting students' final grades.

Socio-demographic data of over 450 students, which were collected during the time of enrollment at the Open Polytechnic of New Zealand were analyzed by Kovacic [14] for predicting students' success. For classification purposes he applied CHAID and CART algorithms and showed that CART transcended CHAID.

Hijazi and Naqvi [15] identified several factors that influenced the students' performance in the intermediate examination using simple linear regression. They found that class attendance, family income, mother's education, and study hours per day have a proportional relation with the student's performance, and mother's age has a reverse proportional relation with the result.

Mia et al. [16], proposed different machine learning techniques for predicting the registration status of the private university's students of Bangladesh. Among the different classifiers, Support Vector Machine outperformed all other classifiers and achieved an accuracy of 85.76%.

Biswas et al. [17] used diverse machine learning classifiers to predict the enrollment and dropout status in the post-graduation level. For this work, they collected the dataset from a renowned public university of Bangladesh. They computed the performance evaluation metrics for each of the classifiers. Finally, they found that the locally weighted learning outstrips the other classifiers.

III. METHODOLOGY

This section is divided into three subsections: data description, algorithms description, implementation procedures. The subsections are briefly described below.

A. Data Description

The dataset used in this study has been obtained from a reputed private university of Bangladesh. It contains the records of 400 students of diverse courses of different departments from Summer 2018 to Fall 2019. This research has been performed using eight attributes, among them only one attribute is the response variable and the other seven attributes are predictor variables. These variables are described below in details.

- ATDM: This attribute depicts the attendance marks of a student.
- RTK: It represents whether a student has retaken the subject or not.
- APAQ: During the tenure of a single semester, a student has to give three quizzes in a particular course. This attribute portrays whether a student appeared in all the quizzes or not.
- AQM: The average of the obtained quiz marks is depicted by this attribute.
- MIDM: The obtained marks in the mid term examination is represented by this attribute.
- SUAS: This attribute confirms whether a student has submitted the assignment or not.

TABLE I. GRADING POLICY OF THE UNIVERSITY GRANTS COMMISSION (UGC) OF BANGLADESH

Marks out of 100	Letter Grade	Grade Point	Marks out of 100	Letter Grade	Grade Point
80-100	A+	4.00	55-59	B-	2.75
75-79	A	3.75	50-54	C+	2.50
70-74	A-	3.50	45-49	C	2.25
65-69	B+	3.25	40-44	D	2.00
60-64	B	3.00	0-39	F	0.00

TABLE II. VARIABLES FOR PERFORMING PREDICTION

Variable Name	Variable Type	Data Type	Possible Values
ATTDM	Predictor Variable	Real Number	0-7
RTK	Predictor Variable	Categorical	Y (Yes) N (No)
APAQ	Predictor Variable	Categorical	Y (Yes) N (No)
AQM	Predictor Variable	Real Number	0-15
MIDM	Predictor Variable	Real Number	0-25
SUAS	Predictor Variable	Categorical	Y (Yes) N (No)
PPRE	Predictor Variable	Categorical	Y (Yes) N (No)
FNLG	Response Variable	Real Number	0 (Fail) 1 (Poor) 2 (Good) 3 (Excellent)

- PPRE: Represents whether a student has performed the presentation or not. The above seven attributes are the predictor variables.
- FNLG: This is the only response variable. It depicts the final grade of a student after the final examination. The university follows the grading policy of University Grants Commission (UGC) of Bangladesh which is shown in Table I [18].

The final grades are categorized into four categories. If a student achieves A+, A or A- , then his grade is categorized into the category ‘Excellent’. B+, B and B- are considered as ‘Good’. The letter grades C+, C, D are categorized as ‘Poor’, and the grade F is considered as ‘Fail’. The possible values, data types, and variable types of different variables used in this research are shown in Table II.

B. Algorithms Description

The algorithms used in this research are described below in details.

1) *Support Vector Machine (SVM)*: Support Vector Machine (SVM) tries to separate two classes using an optimal hyperplane [19]. It uses supervised learning. SVM works better, if the size of the data is small [20]. It attempts to make the decision boundary to such a degree that the partition between two classes is as broad as could reasonably be expected. To separate two classes, let’s assume we are given a training data set, $D = (x_1, C_1), (x_2, C_2), \dots, (x_N, C_N)$ where x_i denotes input vector and C_i refers to the class label of the vector which

could be specified as either positive or negative. For specifying any unspecified vector X , the condition is as follows:

$$f(X) = \sum_{i=1}^N a_i C_i (x_i^T X) + b \quad (1)$$

Here, the nonzero coefficients are a_i ($i = 1, 2, \dots, N$) and the bias is represented by b [21].

2) *Logistic Regression*: The relationship between different variables are settled by regression analysis. If the relationship is linear, then Linear Regression analysis can be applied. But in the case of nonlinear relationship between the variables, we can’t apply Linear Regression and Logistic Regression can be introduced then. Logistic Regression is a generalized form of Linear Regression [22]. Consider the following equation for the Linear Regression:

$$y = \alpha_0 + \alpha_1 Z_1 + \alpha_2 Z_2 + \dots + \alpha_n Z_n \quad (2)$$

Here, y is the response variable and $Z_1, Z_2, Z_3, \dots, Z_n$ are the predictor variables. By applying the sigmoid function on the equation, we can get the logistic function.

$$l = 1/[1 + e^{-(\alpha_0 + \alpha_1 Z_1 + \alpha_2 Z_2 + \dots + \alpha_n Z_n)}] \quad (3)$$

3) *K-Nearest Neighbor (KNN)*: A simple, non-parametric supervised learning algorithm is K-Nearest Neighbor algorithm, which can be used for both regression and classification. Based on the feature similarity (e.g. distance function), all the available cases are stored and new cases are classified by it. The output is a class membership in KNN classification. A case is categorized by a predominance vote of its neighbors.

The case is allotted to the utmost common class among its K nearest neighbors. Various heuristic techniques can select the value of K (positive integer) in KNN method. The case will be assigned to the class of its nearest neighbor if K=1 [23]. Different distance functions like: Minkowski Distance, Manhattan Distance, Euclidean Distance are used in KNN algorithm. In this work, Minkowski Distance function has been used. The Minkowski Distance for two points $U (u_1, u_2, \dots, u_n)$ and $V (v_1, v_2, \dots, v_n)$ can be represented by the following equation, where q represents the order of the Minkowski Distance.

$$distance(U, V) = \left(\sum_{i=1}^n (|u_i - v_i|^q)^{1/q} \right) \quad (4)$$

4) *Decision Tree*: Decision Tree classification uses tree like structures. The internal nodes of the tree represent the conditions and the external nodes or the leaves represent the class labels. The branches from the internal nodes represent the outcomes of the tests or conditions. The decision of splitting the data is controlled by entropy and which can be defined by the equation below, where p_j is the probability of the j^{th} class.

$$E(S) = \sum_{j=1}^c -p_j \log_2 p_j \quad (5)$$

Different variations of Decision Tree are available as for instance: ID3, C4.5, CART, etc.

5) *AdaBoost*: AdaBoost stands for Adaptive Boosting classifier. A set of weak classifiers is combined into a strong one using this approach. Here, the following equation represents the classification using the AdaBoost algorithm.

$$F(t) = sign\left(\sum_{m=1}^M \theta_m f_m(t)\right) \quad (6)$$

Here, the m^{th} weak classifier is represented by f_m and θ_m represents the corresponding weight.

6) *Multilayer Perceptron (MLP)*: Multilayer Perceptron is a form of feedforward neural network and it consists of multiple layers of neurons. A neuron of one layer interacts with the neurons of its adjacent layers through weighted connections though there exists no connection between the neurons of the same layer. Excluding the input and the output layers, the MLP has one or more hidden layers or intermediate layers [24]. The error of the k^{th} output node in the data point n can be represented by the equation below where d and c represent the actual and predicted values respectively.

$$e_k(n) = d_k(n) - c_k(n) \quad (7)$$

7) *Extra Tree Classifier*: A variant of Random Forest known as Extra Tree Classifier was first introduced by Geurts et al. [25]. Extra Tree Classifier differs from other tree based classifiers in such a way that it uses the entire learning sample for growing the trees and it chooses cut-points for splitting the nodes fully at random.

8) *Weighted Voting Classifier*: Voting Classifier is an approach for combining the outputs of different base classifiers as it is hard to identify a specific classification algorithm that gives the best accuracy on a certain data. Both homogeneous and heterogeneous models can be aggregated using the Voting Classifier. In the weighted voting approach, a weight or coefficient is assigned to each base classifier which is proportional to the base classifier's individual accuracy [26]. Consider $h_1, h_2, h_3, \dots, h_n$ are the outputs of n-different classifiers respectively and $s_1, s_2, s_3, \dots, s_n$ are the assigned weights to each classifier, respectively, then the final output H of the Weighted Voting Classifier can be represented by the following equation.

$$H = s_1 * h_1 + s_2 * h_2 + s_3 * h_3 + \dots + s_n * h_n \quad (8)$$

C. Implementation Procedures

The implementation procedures are illustrated in this section. To carry out the study, Python and Scikit-learn library have been used.

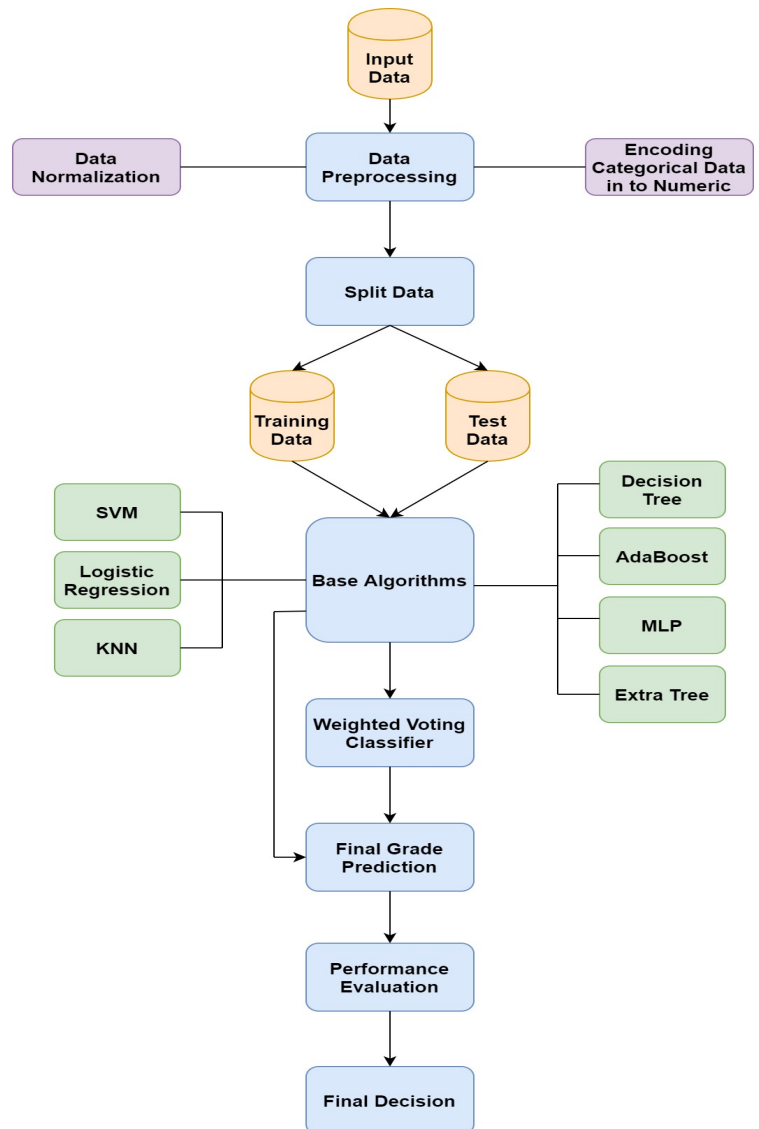


Fig. 1. Step by Step Procedures for Predicting Students' Final Grades

The graphical form of the stepwise procedures for predicting students' final grades is represented by Fig. 1. The details of Fig. 1 is depicted below.

1) *Input Data*: After collecting the data of 400 students via the Enterprise Resource Planning (ERP) system of the university, the task of inputting the data in the proposed system has been performed in this step.

2) *Data Preprocessing*: Data preprocessing step is categorized into two categories, namely: Data normalization and Encoding the categorical data into numeric data. In the collected dataset, the attendance marks range from 0 to 7, the average quiz marks range from 0 to 15 and the obtained mid term examination marks range from 0 to 25. Under the circumstances, these three predictor variables are in very different ranges. So, normalization of these three attributes has been performed. After the normalization procedure, the values of these three variables range from 0 to 1. There are some categorical data in the dataset. Algorithms like Decision Tree algorithm can work effectively with categorical data but most of the other algorithms give better performance while using numerical data instead of its categorical counterpart. Hence, the categorical data have been encoded into numerical data using the Label Encoding approach of the Scikit-learn library.

3) *Data Splitting*: Splitting the dataset follows the data preprocessing step. This step splits the dataset into training

data and test data. In this work, 74% of data is used for training purposes and the rest 26% of data is used for testing.

4) *Training and Testing using Base Classifiers*: In this step, the seven base classifiers have been trained with the training data. And after training the classifiers, the prediction of the final grades of the students has been performed using the test data. Accuracy of each base classifier is also measured separately.

5) *Aggregating the Outputs of the Base Classifiers*: Eventually, using the Weighted Voting Classifier, this step aggregates the outputs of these seven base classifiers for achieving better performance.

6) *Performance Evaluation*: This step compares the performance of the base classifiers with the performance of the Weighted Voting Classifier. For evaluating the performance, five evaluation metrics: Accuracy, Precision, Recall, F-1 score and Area Under Curve (AUC) are determined.

7) *Final Decision*: According to the outcomes of the evaluation metrics, the best classifier for predicting the final grades of the students has been selected in this step.

IV. RESULTS

For testing purposes, the records of 104 students have been used. Among these records, 38% records are actually classified

TABLE III. CONFUSION MATRICES FOR DIFFERENT CLASSIFIERS FOR PREDICTING STUDENTS' GRADES

Classifiers Name	Predicted→	Fail	Poor	Good	Excellent
	Actual↓				
SVM	Fail	2	1	0	0
	Poor	3	29	5	0
	Good	0	4	10	10
	Excellent	0	1	1	38
Logistic Regression	Fail	2	1	0	0
	Poor	2	19	13	3
	Good	0	2	11	11
	Excellent	0	0	3	37
KNN	Fail	3	0	0	0
	Poor	3	30	4	0
	Good	0	5	16	3
	Excellent	0	1	11	28
Decision Tree	Fail	3	0	0	0
	Poor	4	25	8	0
	Good	0	2	19	3
	Excellent	0	1	7	32
AdaBoost	Fail	3	0	0	0
	Poor	3	19	14	1
	Good	0	1	15	8
	Excellent	0	0	4	36
MLP	Fail	2	1	0	0
	Poor	2	24	10	1
	Good	0	2	15	7
	Excellent	0	0	3	37
Extra Tree	Fail	3	0	0	0
	Poor	3	29	5	0
	Good	1	6	13	4
	Excellent	0	1	6	33
Weighted Voting Classifier	Fail	3	0	0	0
	Poor	3	29	5	0
	Good	0	2	19	3
	Excellent	0	0	6	34

TABLE IV. MEASURED RESULTS OF DIFFERENT CLASSIFIERS FOR PREDICTING STUDENTS' GRADES

Classifiers Name	Accuracy	Class Label	Precision	Recall	F-1 Score	AUC
SVM	75.96%	Fail	0.40	0.67	0.50	0.81
		Poor	0.83	0.78	0.81	
		Good	0.63	0.42	0.50	
		Excellent	0.79	0.95	0.86	
Logistic Regression	66.35%	Fail	0.50	0.67	0.57	0.76
		Poor	0.86	0.51	0.64	
		Good	0.41	0.46	0.43	
		Excellent	0.73	0.93	0.81	
KNN	74.04%	Fail	0.50	1.0	0.67	0.85
		Poor	0.83	0.81	0.82	
		Good	0.52	0.67	0.58	
		Excellent	0.90	0.70	0.79	
Decision Tree	75.96%	Fail	0.43	1.0	0.60	0.87
		Poor	0.89	0.68	0.77	
		Good	0.56	0.79	0.66	
		Excellent	0.91	0.80	0.85	
AdaBoost	70.19%	Fail	0.50	1.0	0.67	0.83
		Poor	0.95	0.51	0.67	
		Good	0.45	0.63	0.53	
		Excellent	0.80	0.90	0.85	
MLP	75%	Fail	0.50	0.67	0.57	0.81
		Poor	0.89	0.65	0.75	
		Good	0.54	0.63	0.58	
		Excellent	0.82	0.93	0.87	
Extra Tree	75%	Fail	0.43	1.0	0.60	0.85
		Poor	0.81	0.78	0.79	
		Good	0.54	0.54	0.54	
		Excellent	0.89	0.82	0.86	
Weighted Voting Classifier	81.73%	Fail	0.50	1.0	0.67	0.90
		Poor	0.94	0.78	0.85	
		Good	0.63	0.79	0.70	
		Excellent	0.92	0.85	0.88	

as “Excellent”, 23% records are actually classified as “Good”, 36% records are actually classified as “Poor” and the other 3% records are originally classified as “Fail”.

The confusion matrices of the result of this study using SVM, Logistic Regression, KNN, Decision Tree, AdaBoost, MLP, Extra Tree and Weighted Voting Classifier are presented in Table III.

The calculated Accuracy, Precision, Recall, F-1 Score and AUC (Area Under Curve) are shown in Table IV.

Table IV exhibits that Logistic Regression gives the lowest accuracy of 66.35%, where the Weighted Voting Classifier attains the highest accuracy of 81.73%. After the Weighted Voting Classifier, the second-highest accuracy of 75.96% is attained by the SVM and Decision Tree classifier jointly. The

TABLE V. MEASURED RESULTS OF DIFFERENT CLASSIFIERS USING TWO CLASS LABELS

Classifiers Name	Accuracy	Class Label	Precision	Recall	F-1 Score	AUC
SVM	92.31%	Lower Order Grades	1.0	0.80	0.89	0.90
		Higher Order Grades	0.89	1.0	0.94	
Logistic Regression	81.73%	Lower Order Grades	0.92	0.57	0.71	0.77
		Higher Order Grades	0.78	0.97	0.87	
KNN	90.38%	Lower Order Grades	0.86	0.90	0.88	0.90
		Higher Order Grades	0.94	0.91	0.92	
Decision Tree	83.65%	Lower Order Grades	0.81	0.75	0.78	0.82
		Higher Order Grades	0.85	0.89	0.87	
AdaBoost	83.65%	Lower Order Grades	0.85	0.70	0.77	0.81
		Higher Order Grades	0.83	0.92	0.87	
MLP	84.61%	Lower Order Grades	0.93	0.65	0.76	0.81
		Higher Order Grades	0.82	0.97	0.89	
Extra Tree	77.88%	Lower Order Grades	0.72	0.70	0.71	0.76
		Higher Order Grades	0.82	0.83	0.82	
Weighted Voting Classifier	93.26%	Lower Order Grades	1.00	0.82	0.90	0.91
		Higher Order Grades	0.90	1.00	0.95	

third-highest accuracy of 75% is achieved by MLP and Extra Tree classifier. The accuracy of KNN is 74.04%, while the accuracy of AdaBoost is 70.19%.

The Area Under the Curve (AUC) for different classifiers has been measured also. When the value of AUC for a certain classifier is 1.0 then the classifier is considered as a perfect classifier and if the value of AUC is 0.5, then the classifier is considered as a worthless classifier. Here the achieved AUC value for the Weighted Voting Classifier is 0.90 which has surpassed the AUC value of other base classifiers and the lowest AUC value was achieved by the Logistic Regression. The AUC value of Logistic Regression is 0.76. The AUC values of SVM, KNN, Decision Tree, AdaBoost, MLP, and Extra Tree classifier are 0.81, 0.85, 0.87, 0.83, 0.81, and 0.85 respectively.

Additionally, this study has been performed by reducing the number of class labels also. In this task, the classes 'Excellent' & 'Good' have been categorized as 'Higher Order Grades' and 'Poor' & 'Fail' classes have been categorized as 'Lower Order Grades'. After that the performance of the proposed approach has been measured again. Table V represents the performance metrics using two class labels. From this table, it is found that the Weighted Voting Approach has gained the highest accuracy of 93.26%. Comparing Table IV and Table V, it can be observed that the accuracy has been significantly increased by reducing the number of class labels.

In both cases, the Weighted Voting Classifier has improved the accuracy. So, it can be confirmed that the Weighted Voting Classifier has overshadowed the other classifiers undoubtedly.

This study uses a dataset of 400 students of different courses and different departments. As the study has gathered the data from a variety of departments and with this dataset the proposed approach gained an accuracy of 81.73%, so, it can be assured that the proposed approach is reliable enough.

V. CONCLUSION

This study has used seven base classifiers to predict the students' final grades and then combined the outputs of the base classifiers using weighted voting approach. And from the observation, it can be confirmed that aggregating the base classifiers using the weighted voting approach has caused a rise in the accuracy. From the achieved AUC values it can be also stated that the Weighted Voting Classifier is almost the perfect classifier for classifying the accumulated dataset.

The limitation of this study is, it has not shown any comparison among the performance of this proposed approach and other approaches' performance, illustrating other study.

VI. FUTURE WORKS

This work is performed by using the dataset of only one private university of Bangladesh. In future the dataset can be enlarged by collecting data from different private and public universities of Bangladesh to achieve better performance and better accuracy. Moreover, a comparative study between the proposed approach of this work and the approaches presented in other works can be performed in future.

Different studies show that by preprocessing the data using discretization method and oversampling techniques like Synthetic Minority Over-sampling Technique (SMOTE) can result in an increase of the accuracy. By using these approaches, preprocessing of the gathered dataset can be performed to get better performance.

REFERENCES

- [1] "Private University," Available online: <http://www.ugc-universities.gov.bd/private-universities> [Last Accessed 27 February 2020].
- [2] S. K. Yadav and S. Pal, "Data mining: A prediction for performance improvement of engineering students using classification," arXiv preprint arXiv: 1203.3832, 2012.
- [3] R. Kabra and R. Bichkar, "Performance prediction of engineering students using decision trees," International Journal of computer applications, vol. 36, no. 11, pp. 8–12, 2011.
- [4] A. M. Shahiri, W. Husain et al., "A review on predicting student's performance using data mining techniques," Procedia Computer Science, vol. 72, pp. 414–422, 2015.
- [5] B. Kapur, N. Ahluwalia, and R. Sathiyaraj, "Comparative study on marks prediction using data mining and classification algorithms," International Journal of Advanced Research in Computer Science, vol. 8, no. 3, 2017.
- [6] Z. Liu and X. Zhang, "Prediction and analysis for students' marks based on decision tree algorithm," In 2010 Third International Conference on Intelligent Networks and Intelligent Systems, pp. 338–341. IEEE, 2010.
- [7] M. Sweeney, J. Lester, and H. Rangwala, "Next-term student grade prediction," In 2015 IEEE International Conference on Big Data (Big Data), pp. 970–975. IEEE, 2015.
- [8] S. K. Yadav, B. Bharadwaj, and S. Pal, "Data mining applications: A comparative study for predicting student's performance," arXiv preprint arXiv: 1202.4815, 2012.
- [9] B. Minaei-Bidgoli, D. A. Kashy, G. Kortemeyer, and W. F. Punch, "Predicting student performance: an application of data mining methods with an educational web-based system," In 33rd Annual Frontiers in Education, 2003. FIE 2003., vol. 1, pp. T2A–13. IEEE, 2003.
- [10] Z. Iqbal, J. Qadir, A. N. Mian, and F. Kamiran, "Machine learning based student grade prediction: A case study," arXiv preprint arXiv: 1708.08744, 2017.
- [11] W. L. Gorr, D. Nagin, and J. Szczypula, "Comparative study of artificial neural network and statistical models for predicting student grade point averages," International Journal of Forecasting, vol. 10, no. 1, pp. 17–34, 1994.
- [12] Y. Meier, J. Xu, O. Atan, and M. Van der Schaar, "Predicting grades," IEEE Transactions on Signal Processing, vol. 64, no. 4, pp. 959–972, 2015.
- [13] S. T. Jishan, R. I. Rashu, N. Haque, and R. M. Rahman, "Improving accuracy of students' final grade prediction model using optimal equal width binning and synthetic minority over-sampling technique," Decision Analytics, vol. 2, no. 1, p. 1, 2015.
- [14] Z. Kovacic, "Early prediction of student success: Mining students' enrolment data," 2010.
- [15] S. T. Hijazi and S. Naqvi, "Factors affecting students' performance," Bangladesh e-journal of Sociology, vol. 3, no. 1, 2006.
- [16] Md. Jueal Mia, Al Amin Biswas, Abdus Sattar, Md. Tarek Habib, "Registration Status Prediction of Students Using Machine Learning in the Context of Private University of Bangladesh," International Journal of Innovative Technology and Exploring Engineering(IJITEE), vol. 9, no. 01, pp. 2594-2600, 2019.
- [17] Al Amin. Biswas, Anup Majumder, Md. Jueal Mia, Itisha Nowrin, and Nadia Afrin Ritu, "Predicting the Enrollment and Dropout of Students in the Post-Graduation Degree using Machine Learning Classifier," International Journal of Innovative Technology and Exploring Engineering (IJITEE), vol. 8 no. 11, pp. 3083-3088, September 2019.
- [18] M. H. Bhuyan and S. S. A. Khan, "Motivating students in electrical circuit course," International Journal of Learning and Teaching, vol. 10, no. 2, pp. 137–147, 2018.

- [19] R. Moraes, J. F. Valiati, and W. P. G. Neto, "Document-level sentiment classification: An empirical comparison between svm and ann," *Expert Systems with Applications*, vol. 40, no. 2, pp. 621–633, 2013.
- [20] K. Eashwar, R. Venkatesan, and D. Ganesh, "Student performance prediction using svm," *International Journal of Mechanical Engineering and Technology*, vol. 8, no. 11, pp. 649–662, 2017.
- [21] M. Olgun, A. O. Onarcan, K. Özkan, Ş. Işık, O. Sezer, K. Özgüşi, N. G. Ayter, and Z. B. Başçiftçi, M. Ardiç, and O. Koyuncu, "Wheat grain classification by using dense SIFT features with SVM classifier," *Computers and Electronics in Agriculture*, vol. 122, pp. 185-190, 2016
- [22] A. Dutta, G. Bandopadhyay, and S. Sengupta, "Prediction of stock performance in indian stock market using logistic regression," *International Journal of Business and Information*, vol. 7, no. 1, 2012.
- [23] Y. Y. Aung and M. M. Min, "Hybrid intrusion detection system using k-means and k-nearest neighbors algorithms," In *2018 IEEE/ACIS 17th International Conference on Computer and Information Science (ICIS)*, pp. 34–38. IEEE, 2018.
- [24] S. K. Pal and S. Mitra, "Multilayer perceptron, fuzzy sets, and classification," 1992.
- [25] P. Geurts, D. Ernst, and L. Wehenkel, "Extremely randomized trees," *Machine learning*, vol. 63, no. 1, pp. 3–42, 2006.
- [26] D. Tripathi, D. R. Edla, V. Kuppili, A. Bablani, and R. Dharavath, "Credit scoring model based on weighted voting and cluster based feature selection," *Procedia computer science*, vol. 132, pp. 22–31, 2018.