Academic Advisor: A Prediction of Undergraduates Students Semester Final's Mark with Contextual Feedback Using Machine Learning Approach

Syeda Farjana Shetu
Department of Computer Science and Engineering
Daffodil International University
Dhaka, Bangladesh
shetu4247@diu.edu.bd
https://orcid.org/0000-0002-0691-5127

Fatema Tuj Johora
Department of Computer Science and Engineering
Green University of Bangladesh
Dhaka, Bangladesh
fatema@cse.green.edu.bd

Md. Mehedi Hasan
Department of Computer Science and Engineering
Daffodil International University
Dhaka, Bangladesh
mehedibinhafiz@gmail.com

Marzan Tasnim Oyshi
Faculty of Computer Science
Technische Universität
Dresden, Germany
marzan_tasnim.oyshi@tu-dresden.de
https://orcid.org/0000-0001-9661-1591

Nazmun Nessa Moon
Department of Computer Science and Engineering
Daffodil International University
Dhaka, Bangladesh
moon@daffodilvarsity.edu.bd
https://orcid.org/0000-0001-6641-006X

Abstract—The aim of this research is to predict undergraduate students' academic performance using machine learning techniques. With the increasing availability of instructional data, there is a growing potential to utilize this information for educational purposes. Machine learning has become a common approach to predicting student performance, which can be beneficial for improving teaching strategies and student outcomes. This study focused on identifying challenges faced by graduate students who have low academic performance, and how their future performance can be predicted using historical data. The dataset used in this study was collected from a reputable academic institution and analyzed using various machine learning algorithms, such as Decision Trees, Random Forests, Support Vector Machines, Gradient Boosters, Linear Regressions, and Neural Network Regressions. The most effective algorithm was used to predict students' final semester grades. Feedback and suggestions for improvement were provided to students based on their predicted grades. The proposed system, named Academic Advisor, acts as a coach or guide for students, displaying their current academic status and providing customized targets to help them achieve better grades. This research can help educators and institutions improve their teaching strategies and enhance students' academic performance by utilizing machine learning techniques.

Keywords— education technology, estimation, machine learning

I. INTRODUCTION

There are a lot of factors that determine a student's cumulative grade point average (CGPA), including prior course performance and credit earned for subject study in the past. Every student aspires to maintain a grade point average that is as high as reasonably achievable. When a student's cumulative grade point average (CGPA) is calculated, it represents their overall level of academic success. It is necessary for students to examine a range of elements in order to maintain a high cumulative grade point average (CGPA) throughout their academic careers. Attendance, presentation,

and concentration in class are all important factors in achieving success; just studying will not be sufficient on its own. Finally, the primary goal of this investigation is to assess the significance of each of these factors in order to get the best possible outcome for the participants. To find the most important factor in achieving a high cumulative grade point average, machine learning may be employed (CGPA). The primary goal of this study is to predict a student's final test score based on his or her previous performance and to assess whether or not the student is at risk for learning problems in the future. This topic is straightforward enough that students may easily identify their own weaknesses while also gaining a good estimate of their final grade on it. This strategy also includes suggestions and alternative ideas for increasing their overall performance. Students will have an easier time preparing for the final test as a result of using this method of learning. Providing that a student is able to overcome a weakness, it may be assumed that more predictable outcomes will be attained in the future. Once the data had been obtained, it was merged and recognized using a variety of machine learning techniques that were developed. The dataset is at the heart of all machine learning algorithms since it contains all of the information. The grades of pupils were utilized to compile the data for this inquiry, which was then analyzed. Daffodil International University in Bangladesh has given this information, which contains the grades of its students, for your consideration. Aside from the development of features and labels, it was able to immediately categorize the data. Quizzes, assignments, presentations, midterms, and attendance were all taken into consideration while compiling the data set. How can they determine the limitations of children's abilities based on their examination results?" For everything from semester numbers to examinations and assignments to presentations, midterms, attendance, and everything in between, Researchers established a standard label that could be applied to everything. They used it for everything. The use of machine learning to forecast the future has a wide range of repercussions. The plan is to use the Machine Learning technique to forecast the future results of students based on the

dataset in order to boost the efficacy and understanding of the system while keeping the cost as low as possible. In its development and research, they have employed a variety of machine-learning packages, including: Scikit-Learn, Pandas, Numpy, and Matplotlib, to name just a few examples. In order to construct an adequate dataset, it is required to use a number of different feature selection techniques. The field of Machine Learning encompasses many various sorts of predictions, including enhanced algorithms, supervised algorithms, and uncontrolled algorithms, amongst others. Therefore, regression analysis is used in order to predict a student's final academic achievement. The method accurately predicts students' academic achievement, and one of its main goals is to assist students in improving the academic performance in the classroom. This is accomplished by developing customized suggestions for each individual student based on the student's abilities and mental process, among other factors.

II. LITERATURE REVIEW

Machine learning is extensively employed for determining matters that are obedient to forecasting. For exerting steps against the student's future performance prediction, a lot of work has been made using ML. ML has performed this procedure much conveniently. B. Minaei-Bidgoli, et al. implemented data mining techniques on the dataset, which is obtained from the LONCAPA database. Mainly they worked on prognosticating student performance to help web-based educational technology. By this work, the authors encouraged the student to find and classifying their problems. In their work, Quadratic Bayesian, INN, KNN, Parzen-window, and decision tree, Classifiers are used. By joining the classifier with multiple classifiers, they develop classifier performance. They mapped the GA with CMC for improving the result. For optimizing CMC with GA they got the best accuracy [1]. S. B. Kotsiantis, et al. have done their work on the purpose of predicting student performance by their results. The central technique of their work was the application of a regression algorithm to attain student performance whether a student passes a course or not. Finally based on their work they built a software prototype for this purpose. They used ML techniques to find out the poor performance of a student to notify them. HOU gave them the required database from it they extracted the dataset for further proceedings. The feature was divided into three categories. They used common regression methods such as model trees, neural networks, linear regression, locally weighted linear regression, SVM. They used data collection. From model tree inductor M5 which works on propositional regression rule helped them to acquire the best accuracy. Finally based on the M5 rule result they constructed software support tools [2]. M. Ross, et al. done this work to classify student attentiveness. They found whether a student is attentive or not by the ML approach. They used K-means clustering and SVM for prediction. For data collection, an RGB-D sensor was used. It stored student's gestures, postures, and facial expressions. Extracting from its data was produced to find out student attentiveness and algorithm implementation. The algorithm was implemented for classifying the student's attentiveness or inattentiveness. After making the required dataset it was clustered by K-means and classified by the SVM algorithm. Finally with the help of this ML technique they built a system for identifying student attentiveness automatically [3]. H. Hamsa, et al. have done similar work to enhance the quality of higher education. The authors collected data from 120 bachelor's students, 48 master's students. They examined the dataset into training and testing data. Admission score, sessional marks, internal marks were considered as attributes. They utilized Decision Trees (DT) and the Fuzzy Genetic Algorithm (FGA) for prediction. Based on the result they build a model. Finally, they showed different results to the respective (instructor, student) entities for their better actions [4]. M. M. Mohan, et al. showed that amount of educational data is rising tremendously. They worked to find the hidden structure that lies in the data by big data analysis. Based on this they foretell the performance of students. They used big data techniques like Hadoop, Map Reduce for this analysis [5]. A. Acharya, et al. discussed EDM on their work. For the early prediction of student performance EDM applications were employed. To offer remediation to the week students, it is important to predict their performance early. For this work, they derived their dataset by collecting data from the colleges of Kolkata. For building a prediction model, 15 attributes are considered. After processing the data, they used ML algorithms on the dataset including Decision Tree, Bayesian Network, ANN, and SVM. They analyzed the accuracy of the model with the help of the confusion matrix. They found SMO and C4.5 algorithms work fit on their dataset [6]. B. Khan, et al. showed that the database holds information about students from which hidden patterns can be observed. Data were obtained from the student of S.S.C of Islamabad. They divided the main method into learning and classification. After processing the data, algorithms are implemented for correct classification and prediction. They applied the J48 decision tree algorithm, which is an implementation of the decision tree by JAVA. With the implementation of this algorithm, they develop a model to predict student's final grades. They got 84.53% accuracy by their model. Their prediction helps students, parents, and teachers to take advanced initiatives to enhance the performance of the student. After going through all the literature discussed preceding, at contrast to their research, they have discovered that they also have several common peculiarities and variation. In this work, fourteen thousand data which is huge and predicts the student's final mark also give proper suggestions for individual students to improve their performance in university is discussed. They used four traditional ML algorithms to achieve the best outcome [7]. Adekitan et al (2019) despite the fact that academic and nonacademic elements impact college or university achievement. For example, high-achieving students may lose concentration owing to peer pressure and social diversions, whereas lowachieving students may succeed in university. The same cannot be said about academic excellence in Nigeria. It was determined that the link between cognitive entrance criteria and first-year academic achievement was linear. The R2 values of 0.207 and 0.232 show that the cognitive entry criteria do not fully predict first-year student performance [8]. Osmanbegovic et al (2012) using preoperative assessment data, three supervised data mining methods were evaluated for accuracy, learning easiness, and user-friendliness. Outperforms decision trees and neural networks. A good classifier model is exact and clear. This was done following a standard classroom data collection. This strategy may help students and teachers improve grades and reduce failures. Interaction influences pleasure. Three supervised data mining techniques performed well. DTs and NNs are outclassed. A good classifier model is accurate and easy to train. Then came data mining in the classroom. This strategy may help students and teachers. Interaction affects both engagement and productivity and student results [9]. Saleh et al (2021) the latest current discoveries in EDM

research are discussed in this paper. There was a lot of focus on the educational aims of each study and on the data and data mining methods used. It's time for a new classroom! When it comes to EDM, elearning, data mining, and tutoring systems are all intertwined with one another. EDM The fast evolution of educational data analysis may be seen through the ever-expanding volume of data. Data mining in education has just recently emerged. This region's future is exciting to contemplate. Data mining is used to predict student outcomes in this research [10]. Ramdas et al (2019) despite the fact that academic success depends on prior performance. Prior performance affects student success. Bigger data equals better SVM Algorithm tracks student's academic and extracurricular activities. Manual examinations were recently employed. Flaws in the heel slow currents Manual exam analysis is challenging. Hand computations are inaccurate. This process is slower. Pre-school test Portal is born. Not an easy task. Designed for teachers and students. Students' academic progress may be predicted. It employs neural networks. The value of some assets is also assessed [11]. Adekitan et al (2019) describe the advantages of machine learning are increasing educational data mining study. The Konstanz Information Miner was used to forecast the fifth year and final Cumulative Grade Point Average (CGPA) of Nigerian engineering students (KNIME). Six data mining algorithms were tested with 89.15 percent accuracy. Linear and pure quadratic regression models had R2 values of 0.955 and 0.957. This allows for early detection of individuals who may not graduate or who may not graduate satisfactorily [12]. Geetha et al (2021) states about that massive data storage has always been an issue. The amount of instructional data expands as awareness develops. This requires a new machine learning approach. Predicting student performance may help administrators, educators, and students avoid student failure. This may also assist pupils improve next semester. They used XGboost, KNN, and SVM to develop prediction models. To discover the most accurate approach, it analyzed accuracy, precision, and recall. SVM and K-NN beat XGBoost in predicting underperformers [13]. Shingari et al (2018) create a mining academic dataset for patterns relevant to administrators, teachers, and students. Education evolves, and students must adapt. This article concerns data mining student records. The best way to predict a student's ultimate grade before acting. It analyzed a group of students' academic records utilizing data from a famous university. Various data mining techniques were used to establish unique categories. If educators use this method early, they may be able to help students in need sooner [14]. Felix et al (2020) learning to code is one of the most difficult tasks for computer science students. Students who struggle to master new skills (like abstraction) are more likely to drop out of classes. To avoid or cultivate specific habits or tendencies, a student should learn the aspects that contribute to success or failure early in their academic career. A computer programming class was given four questionnaires. Machine learning predicts students' grades based on survey responses. Using the data, they can forecast how many kids will require extra help. Motivated folks are more interested in the topic. So, it can reduce dropout rates and improve the bar for all kids [15]. Altabrawee et al (2019) aims to provide research about students with advanced degrees have a right to demand the best. To achieve this goal, greater help for such students is required. Al-Muthanna University's College of Humanities used four machine learning algorithms to predict student achievement in a computer science course. They utilize A.N.N. and decision trees. Students' time on Facebook and the utilization of social media as a learning aid were examined. Students' usage of social media and online education. The accuracy of each model's classification was assessed. It used student questionnaires and also grade books. to determine categorization error, memory, and F. On the other hand, the ANN model's classification accuracy is (fully linked feed forward multilayer ANN). The decision-tree method discovered five factors of student achievement [16]. Sujatha, G. et al (2018) graduate students deserve the finest. These kids need extra help. Al-AI Muthanna's forecasts student improvement. ANNs and conifers these were examined. So did smartphones. With so many applications, cellphones have grown in popularity. They must now concentrate on customer-specific solutions. Determine mobile software preferences based on age, gender, and activity. Authors used SVM and ANN to predict the type of app. Real-time mobile app prediction is possible. Their approaches work. On Facebook to learn the student's internet use I tried it out. A+ for classification accuracy (fully linked feed forward multilayer ANN). DT predicts student achievement [17]. Xu et al (2017) despite that to graduate on time, students must be able to forecast future achievement. It is possible to forecast a student's success on exams and in solving problems, but not their degree completion (e.g., college programs). Some courses predict outcomes better than others. The predicted growth in student numbers must be considered. Machine learning may be used to predict a student's degree program success. Author's suggestions a two-tiered predictor cascade is used. This method uses probabilistic matrices and latent component models. The proposed strategy beat the benchmark procedures in a three-year UCLA research [18]. Muñoz-Bullón et al. (2017) an extracurricular sport's influence on kids' academic progress was the study's purpose. Prior research had conflicting results, with some seeing a positive benefit and others a negative. They aim to offer a more comprehensive view of the outcomes. The empirical data come from a 2008–2014 Spanish university undergraduate student panel. The academic performance of athletes is compared to non-athletes. Taking part in organized sports correlates with higher grades. The research confirms that athletic activities assist practitioners meet academic achievement goals [19]. Shetu et al [20] also predicted the students' performance which is pretty much similar ideology as mine.

III. PROBLEM STATEMENT

University students in our country are frequently apathetic about evaluating themselves according to their academic assessment standards, and it is evident that they are unable to get good results in the end due to the absence of effective preparation and strategy. In our system, Students may estimate the semester's final grades using past data such as attendance, class tests, assignments, presentations, and midterm grades. A student can figure out his or her targeted marks for an exceptional score by doing so. It also has a downside, the predicted marks may not always be sufficient to achieve a student's desired result; in this case, the student can improve their midterm marks by retaking the midterm exam, and after improving their midterm marks, they can aim for moderate final marks to achieve an excellent result. A student can also forecast his or her ultimate outcomes before beginning a semester by using our technique. He or she can create many draft marks on their assessment criteria points and try multiple times to achieve the ideal final result. After finding his/her

draft marks for an excellent result, he/she can smoothly pass the semester.

IV. PROPOSED SYSTEM

The approach includes an absolute of six measures that conclude the analysis, which is shown in Figure. 1.
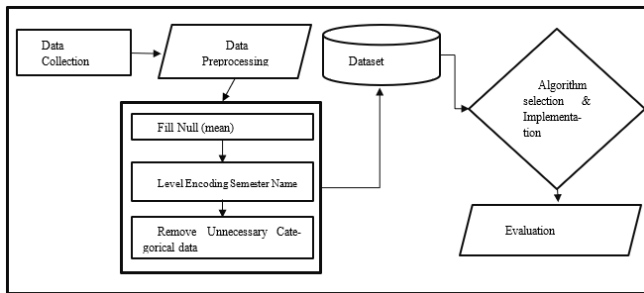

Fig. 1. Methodology Diagram

A. Data Acquisition / Data description

Data compilation is constantly a tough responsibility for all research. Dataset is collected from Daffodil International University in Bangladesh. In this dataset, there are seven attributes such as student semester number, assignments, presentations, Class tests, attendance, midterm test, and final examination marks. This student marking system has been operating at DIU. The complete hundred mark for a student remained separated into six categories that are exhibited in Table 1. Collected all subject's data which marking is matched following Table 1.

TABLE I.   MARKS DISTRIBUTION

Attendance	Class Test	Assignment	Presentation	Mid Term	Final Term
7	5	5	8	5	0

B. Dataset

They demanded to accumulate 14,000 students mark for this research. After assembling 14,000 student's marks, data was pre-processed to build a proper dataset for research. All the data were properly organized into a CSV file. By proper analysis, some unusual information (name, mail, etc.) was removed from the dataset to make e adjustable dataset for the work.
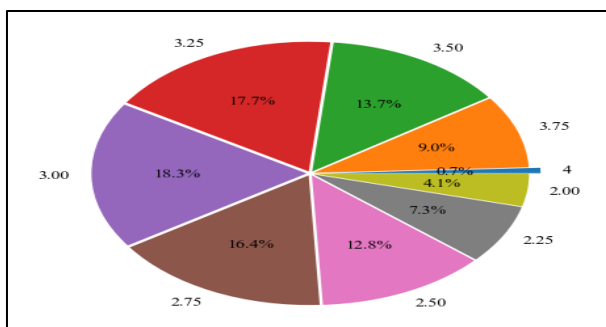

Fig. 2. Database Representation

Figure 2. Represents the overall percentage of the grade of dataset. The figure illustrated that SGPA 3.00 out of 4 is occupied highest amount that is 18.3% and SGPA 4, that is occupied only 0.7% among 14000 students.

C. Data Analysis

After creating a dataset, it is remarkably essential to analyze this data to recognize which algorithm is perfect for this dataset. After analyzing the data so carefully, it found some associations for different attributes with each other, also found out the dependent and independent variables or attributes in the database, so for implementing algorithms, it is easy to decide on what will happen and how it works. Table 2 showed that the independent and dependent classes. Here 'final' attribute is dependent, and the others attribute are independent. It is so crucial to analyze before employing algorithms.

TABLE II.   DEPENDENT & INDEPENDENT VARIABLES

Semester	Attendance	CT	Assignment	Presentation	Mid Term	Final Term
2	0	4	0	1	11.50	28.50
5	3	7.5	0	2.25	16.50	22
9	4	10	0	7	14.50	27
3	6	12	4	7	20.50	27
3	6	10	3	6	16	20

D. Algorithm Selection and Implementation

In the work, the algorithms were concentrated by which are the most suitable in relation to the model. Six conventional machine learning regression algorithms such as Neural Network, Decision tree, Random Forest, and SVM, Linear Regression were employed. Gradient Boosting to get primary efficiency. By this process, it discovered the best algorithm. With the leading highest score among all the algorithms. After executing algorithms, Linear Regression gave the most leading 0.99 r2 scores by employing a 70% training rate. The additional five algorithms also accomplished quite fit. As Linear regression provided the peak efficiency. It has been decided to employ this algorithm to forecast the student result. Table 3 represents the parameter usage of applied algorithms. They selected the parameter that is produced the best performance.

TABLE III.   PARAMETER USAGE

Algorithms	Details
Linear Regression	$Random_state = 3$
Neural Network	$Random_state = 3,$ $max_iter = 500$
Decision Tree	$Random_state = 42$
Support Vector Machine	$kernal =' rbf', degree = 3, gamma = 'scale'$
Random Forest	$n_estimatos = 100, criterion = 'mse'$
Gradient Boosting	$n_estimators = 200$

E. Evaluation

In this work, the system receives input like assignment marks, attendance, class test, presentation, mid exam mark from the student for a particular subject/course. After that, the system interprets the mark and furnishes the Final result as output for that subject additionally it also provides contextual feedback to enhance the student's final result.

V. MAIN EXPERIMENTAL RESULT AND DISCUSSION

To evaluate the effectiveness of its work, six machine learning algorithms were employed in pre-organized in this data. The performance value delivered by these algorithms into Score Matrix Table 4, so it can simply comprehend and implement an association among them based on their attainment. 30% test and 70% training data to estimate the effectiveness of those chosen algorithms. By associating those six algorithms, a magnificent outcome was found. In this work, all the algorithms performed better. Linear Regression performed perfectly with high-level efficiency than other algorithms.

TABLE IV.   ALGORITHM'S SCORE

Algorithms						
Parameter	NN	DT	RF	SVM	LR	GB
MAE	0.16	0.53	0.29	0.24	0.16	0.32
MSE	0.04	0.70	0.20	0.15	0.03	0.18
RMSE	0.20	0.84	0.45	0.39	0.13	0.43
R2 Score	0.985	0.974	0.982	0.987	0.999	0.993

Table 4 represents the different Scores of each algorithm. For each algorithm, their performance measured by Mean Absolute Error (MAE), Mean Squared Error (MSE), Root Means Squared Error (RMSE), and R2Score were evaluated. All of the algorithms generated MSE scores near 0 and r2Score very near to 1. That means each algorithm fitted that data outstandingly. V. Vapnik [21] demonstrated that in ML and pattern distribution SVM gained more stable accomplishment in consequence of the minimization of perceived data. In this research, the SVM algorithm obtained 0.987 r2sore and MSE is 0.15. Linear regression produced the best performance among all the algorithms. C. Jin et al [22] told that the Decision tree is one of the essential procedures for handling model classification and regression for induction analysis and data mining. The Decision tree algorithm gave r2 score 0.974 but 0.70 MSE score and it the highest MSE score among all the algorithms. P. O. Gislason et al [23] described that the Random Forests can retail high dimensional data and manage a huge number of trees considered for classification of multisource. The algorithm Random Forest gave 0.982 r2score, but 0.20 MSE and MSE is suitable. The Neural Network (NN) performed similarly with the Linear Regression. Kangarani Farahani et al [24] stressed that NN is primarily used dynamically for complex prediction and NN more in harmony with denouncing results. 0.982 r2score and 0.04 MSE, which is similar to the linear regression algorithm, were provided by Neural Network. It has achieved the second-best score of all algorithms. Linear regression is the most common statistical model for evaluating the relationship between the various variables. The definition is linear, apart from univariate or multivariate data forms [25]. It may also be a linear regression simple and multiple. Multi-linear regression was used in this work. The highest efficiency was achieved by linear regression with r2score 0.999 and 0.03 MSE using 30% test data rate and 70% training data which is presented in Table V by a red rectangular border-box. The tree method of gradient boosting integrates additional trees strategically by correcting errors of its predecessor versions, thereby improving prediction precision [26]. It observed in their analysis, 0.993 R2 Score and 0.18 MSE are obtained in the gradient boosting algorithm. And its's r2 score is the second highest among all the algorithms. From the above comparative discussion, it can see that linear regression produced the better performance among all the algorithms. So Linear Regression for Evaluation and implementation was decided to use Figure 3. Demonstrated the real and predicted marks of students. The red line represents the real marks, and the green line shows the predicted mark. The green line is so connected to the red line, and it almost overlapped the real mark (red line). That indicates selected model prognosticates very precisely in final marks prediction. From the above discussion, they can determine that the algorithms applied in this research worked sensibly fit and the selected model works quite well. This shows their work excellence.

A. Demo Input Representation

They attempted to develop an intelligence website by the Django framework based on this work. The representation of the website is proffered below. Figure 4 describes the input section of the system. Here, the system gets input Quiz mark, assignment marks, attendance mark, presentation mark, and mid exam mark of the student for a distinct subject/course. Then the system is required to enter or click the Evaluate button to execute the algorithm for the final mark
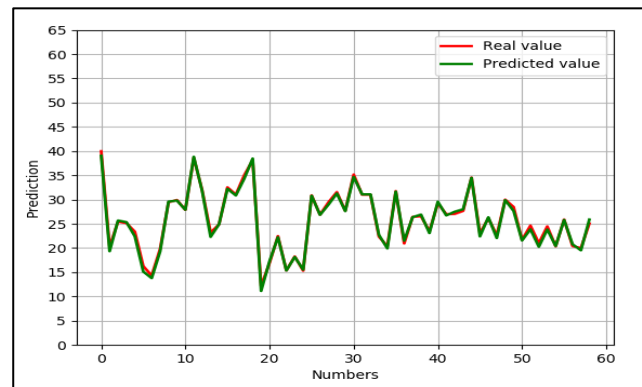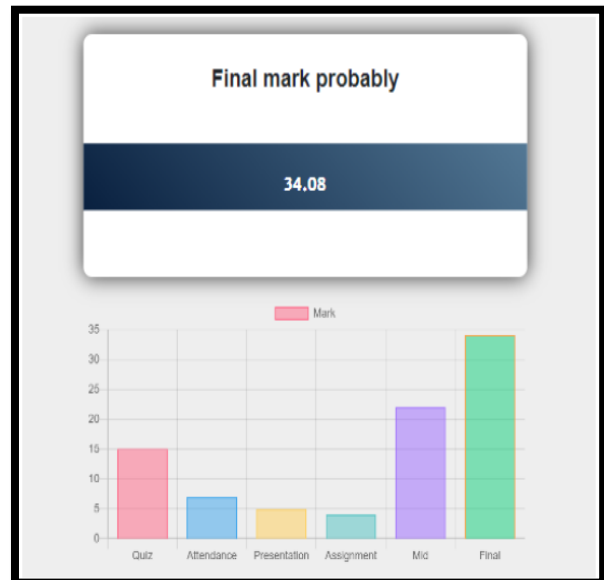

Fig. 3.   Real and Predicted Comparison

Fig. 4.   Input Demonstration

B. Demo Output Representation

Figure 5 depicts the output portion of the system. Based on the input mark algorithm provides the final mark of a student. This system/website is developing stage.

Fig. 5. Output Demonstration

C. Contextual Feedback

To observe the weakness of a student they have confronted his/her all marks with standard marks The standard marks Table 5 is bellowed: If any student got bellow of this standard marks. Then the arrangements for counseling have been made for this student to get more satisfying marks in the final examination. In the credit system of the result, the maximum result is 4.00 and the minimum pass result is 2.00.

TABLE V. STANDARD MARK

Attendance	Class Test	Assignment	Presentation	Mid Term
75% of total class	11	3.5	6	15

By applied this result system, nine groups of students based on CGPA were build and found. Then the corresponding type of 20 students for each group and finally create a total of nine groups were consolidated. After that, the same questionnaires for all groups were applied. The questionnaires are given below:

Questionnaire:

- What is the consequence of each group of students if they are not attending at least 75% of the total class?

- What type of inspection is necessitated if class test marks are bellow standard marks for a particular group?

- How to become a smart worker if assignment marks are bellow standard marks for a specific group?

- How to enhance English speaking, body language, and make wonderful presentation slides if presentation marks bellow standard marks for any group?

After that, some differences, and correlations between all the groups were found. By this comparison, the static approach is not appropriate for all was apprehended. Because different students carry different comprehension and quality. So, if a suggestion for students based on the same mentality and quality will produce, then the suggestion is perfect and accurate for the students. By following questionnaires, the type of process, activities, ability to solve, process overcome, and strategy of each questionnaire was produced. The database into good form with a huge number of data was preprocessed. After the analysis and calculations, it acts as a suggestion making database. By giving this appropriate suggestion to students, they can improve their results by following this.

Algorithm 1: The algorithm to give suggestion.
1: procedure the Procedure
2: top:
3: mark ← take a mark of single attribute
4: loop:
5: if mark < standard mark of specific attribute then
6: goto suggestion.
7: else
8: goto top.
9: end if
10: suggestion:
11: quality ← based on taken mark.
12: if quality == low then.
13: provide low based suggestions.
14: else if quality == mid then.
15: provide medium based suggestions
16: else.
17: provide expert-based suggestions
18: end if
19: end procedure

Figure 6 illustrated contextual feedback to improve the student's final result based on the weakness of the student by executing the aforementioned algorithm. This figure exhibits the suggestion of that student. The suggestion is provided based on students' categories. For example, expert-level suggestions are never provided to the medium or low-level student.

Fig. 6. Provided Suggestion

VI. CONCLUSION AND FUTURE WORK

Following the proposed model, six machine learning algorithms were implemented to predict students' final marks before they attend the final examination and give them appropriate suggestions for individual students. By getting the predicted final mark and suggestions they can enhance their final mark to improve their CGPA. In this proposed model Linear Regression obtained the highest score with 0.999. The linear regression model predicted so precisely that it nearly overlapped the actual result. They ventured to perform the best possible consequence but there are still a few barriers to work. The main limitation of this work is that model works only for students at Daffodil International University (DIU) and universities that follow the same marking system as DIU maintains. In the future, they will try to collect data from all universities in Bangladesh and try to make an intelligent web-based system. So, every university's student can see their predicted results and get their suggestions and guidelines to

improve their performance by giving their past marks and required fields. This intelligent website is in the developing stage, and they are hoping that if proposed system can be developed quickly, this will be very helpful in the corona pandemic situation for academic result processing. In Summary, this a developing prototype & still it is ongoing and soon we will be completed the system on live.

REFERENCES

[1] Minaei-Bidgoli, B., Kashy, D. A., Kortemeyer, G., & Punch, W. F. (n.d.). Predicting student performance: an application of data mining methods with an educational web-based system. 33rd Annual Frontiers in Education, 2003. FIE 2003. doi:10.1109/fie.2003.1263284

[2] Kotsiantis, S. B., & Pintelas, P. E. (2005). Predicting students marks in Hellenic Open University. Fifth IEEE International Conference on Advanced Learning Technologies (ICALT'05). doi:10.1109/icalt.2005.223

[3] Ross, M., Graves, C. A., Campbell, J. W., & Kim, J. H. (2013). Using Support Vector Machines to Classify Student Attentiveness for the Development of Personalized Learning Systems. 2013 12th International Conference on Machine Learning and Applications. doi:10.1109/icmla.2013.66

[4] Hamsa, H., Indiradevi, S. and Kizhakkethottam, J.J.: Student academic performance prediction model using decision tree and fuzzy genetic algorithm. In: Procedia Technology, 25:326–332, 2016.

[5] Mohan, M. G. M., Augustin, S. K., & Roshni, V. S. K. (2015). A BigData approach for classification and prediction of student result using MapReduce. 2015 IEEE Recent Advances in Intelligent Computational Systems (RAICS). doi:10.1109/raics.2015.7488404

[6] Acharya, A., and Sinha, D.: Early prediction of student's performance using machine learning techniques. In: International Journal of Computer Applications, 107(1), 2014.

[7] Khan, B., Khiyal, M.S.H., and Khattak, M.D.: Final grade prediction of secondary school student using decision tree. International Journal of Computer Applications, 115(21), 2015.

[8] Adekitan, A. I., & Noma-Osaghae, E. (2019). Data mining approach to predicting the performance of first year student in a university using the admission requirements. Education and Information Technologies, 24(2), 1527-1543.

[9] Osmanbegovic, E., & Suljic, M. (2012). Data mining approach for predicting student performance. Economic Review: Journal of Economics and Business, 10(1), 3-12.

[10] Saleh, M. A., Palaniappan, S., & Abdalla, N. A. A. (2021). Education is An Overview of Data Mining and The Ability to Predict the Performance of Students. Edukasi, 15(1).

[11] Ramdas, B. R., Machhindra, M. B., Kailas, W. P., & Raut, M. V. (2019). Tracking and Predicting Student Performance Using Machine Learning.

[12] Adekitan, A. I., & Salau, O. (2019). The impact of engineering students' performance in the first three years on their graduation result using educational data mining. Heliyon, 5(2), e01250.

[13] Geetha, R., Padmavathy, T., & Anitha, R. (2021). Prediction of the academic performance of slow learners using efficient machine learning algorithm. Advances in Computational Intelligence, 1(4), 1-12.

[14] Shingari, I., & Kumar, D. (2018). Predicting Student Performance Using Classification Data Mining Techniques. International Journal of Computer Sciences and Engineering, 43-48.

[15] Felix, C., & Sobral, S. R. (2020, April). Predicting students' performance using survey data. In 2020 IEEE Global Engineering Education Conference (EDUCON) (pp. 1017-1023). IEEE.

[16] Altabrawee, H., Ali, O. A. J., & Ajmi, S. Q. (2019). Predicting students' performance using machine learning techniques. JOURNAL OF UNIVERSITY OF BABYLON for pure and applied sciences, 27(1), 194-205.

[17] Sujatha, G., Sindhu. S., & Savaridassan, P. (2018). Predicting students performance using personalized analytics. International Journal of Pure and Applied Mathematics, 119(12), 229-238.

[18] Xu, J., Moon, K. H., & Van Der Schaar, M. (2017). A machine learning approach for tracking and predicting student performance in degree programs. IEEE Journal of Selected Topics in Signal Processing, 11(5), 742-753.

[19] Muñoz-Bullón, F., Sanchez-Bueno, M. J., & Vos-Saz, A. (2017). The influence of sports participation on academic performance among students in higher education. Sport Management Review, 20(4), 365-378

[20] Shetu, S. F., Saifuzzaman, M., Moon, N. N., Sultana, S., & Yousuf, R. (2020). Student's Performance Prediction Using Data Mining Technique Depending on Overall Academic Status and Environmental Attributes. 3rd International Conference on Innovative Computing & Communications (ICICC) 2020, Advances in Intelligent Systems and Computing, 757–769. https://doi.org/10.1007/978-981-15-5148-2_66 Indexing: ISI Proceedings, EI-Compendex, DBLP, SCOPUS, Google Scholar, and Springerlink

[21] Vapnik, V.: The nature of statistical learning theory. In: Springer science & business media, 2013.

[22] Jin, C., De-lin, L., & Fen-xiang, M. (2009). An improved ID3 decision tree algorithm. 2009 4th International Conference on Computer Science & Education, 127-130.

[23] Gislason, P.O., Benediktsson, J.A. and Sveinsson, J.R.: Random forests for land cover classification. In: Pattern recognition letters, 27(4):294–300, 2006.

[24] KangaraniFarahani, M., & Mehralian, S. (2013). Comparison between Artificial Neural Network and neuro-fuzzy for gold price prediction. 2013 13th Iranian Conference on Fuzzy Systems (IFSC). doi:10.1109/ifsc.2013.6675635

[25] K., S., V., S., & R., R. (2016). A comparative analysis on linear regression and support vector regression. 2016 Online International Conference on Green Engineering and Technologies (IC-GET), 1-5.

[26] Zhang, Y., & Haghani, A. (2015). A gradient boosting method to improve travel time prediction. Transportation Research Part C: Emerging Technologies, 58, 308–324. doi:10.1016/j.trc.2015.02.019