MJSAT

Malaysian Journal of Science and Advanced Technology

journal homepage: https://mjsat.com.my/

Grading System Prediction of Educational Performance Analysis using Data Mining Approach

Mahfujur Rahman*[1], Mehedi Hasan[1], Md Masum Billah[1], and Rukaiya Jahan Sajuti[2]

[1] *Dept. of Computer Science, American International University-Bangladesh. Dhaka, Bangladesh.*
[2] *Department of Management Studies, Jagannath University, Dhaka, Bangladesh.*

KEYWORDS

Text Classification
Data Mining
Machine Learning
Predictive Model
Educational Development

ARTICLE HISTORY

Received 7 October 2022
Accepted 9 November 2022
Available online 16 November 2022

ABSTRACT

In the neoteric century, education holds the key to bringing tremendous upgradation to the world. In most Asian countries, it is very challenging to apply education data mining techniques due to the variety of institutional data categories. In this research, an efficient data collection technique has been designed to gather institutional data, analyse and pre-process the data and apply specific data mining methods to estimate students' progress. A real-time dataset has been designed from student transcript data, which helps to analyse the prediction of student quality. In our research, six traditional classification algorithms and a deep neural network (DNN) model is applied to perform prediction efficiency. Different classification models perform an accuracy of 90% ~ 94%. Our research predicts student education efficiency, analyses student patterns and introduces a generalized framework for an advanced level of study.

1. INTRODUCTION

The primary concern of an institution is to provide a quality education that helps to improve student performance. Researchers are constantly on the verge of exploring the factors that emphasize educational growth. They are working on increasing the quality of education, bringing positive changes to the educational system, and enhancing student performance. Recently, educational institutions attempt to find the factors that have a significant impact on gathering the latest technological wisdom for their students. Researchers are constantly bringing up techniques such as data mining to deal with large-scale data generated by educational institutions quickly. Data Mining is a technique that analyzes and extracts information that identifies concealed patterns in the data [1]. It analyses large datasets and discovers patterns and relationships, then used to identify a possible solution. Nowadays, educational institutions have used higher-quality data mining methods that provide quality education. Education performance prediction research is based on data mining skills and strategies that properly collect the high potential data and help predict academic outcomes accurately [2]. When we use actual data to perform the data mining process, it ensures generate accurate results for predicting institutional performance. So educational data mining plays an essential role in developing a nation's potential in the educational sector.

The progressive emphasis on the education data mining system has formerly made a modern experimental method called Educational Data Mining (EDM), which is outbound and concerned with thriving methods for analyzing educational data to understand learners better [3-4]. It is attached to various techniques like decision trees, k-nearest neighbors, association rules, neural networks, genetic algorithms, exploratory factor analysis, and step-wise regression. In the majority of the countries in Asia, various rhythms are noticeable to a student when their performance is slow or down into their undergrad stage [5]. As a result, it generates a massive problem for both faculties and learners to improve institutional performance and take fundamental moves to eliminate the awful outcomes. Predicting academic outcomes and identifying patterns of failure assist teachers in effectively guiding their energetic students by providing upgrade-level instruction and providing extra attention to those who fall below the average line.

*Corresponding author:
E-mail address: Mahfujur Rahman <mahfuj@aiub.edu>.
2785-8901/

In the institutes of Asian countries, we can observe that there has no linear relationship between undergraduate student performance with their previous academic records [6]. Suppose, we divide all of the students into three major categories. In that case, we can see that one category that can understand the topic clearly in the beginning and consistently perform well while the second category reveals the reverse occurrences and be serious in the ending consistently achieve a poor score. And, a few students, the third category, always perform either better or worse. These various types of cases produce difficult for instructors to detect those students who are in risky situations. Most of the students in the last category don't attend classes, can't understand their lessons properly, and then misses their evaluations. If the instructors can find the actual progress of these students, they easily can ensure proper guidance to reduce their dropouts. Some of the students after completing all the subjects of their undergraduate program also turn into major problems and miss their graduation convocation for obtaining poor CGPA (Cumulative Grade Point Average) in all previous semesters. They become very frustrated with their situation and gradually fall apart from their study. If students can guess their situation before falling into big trouble, they can focus on their studies enthusiastically and can improve themselves. If they are dedicated to changing their quality, performance prediction will support them in eliminating the dropouts.

The prediction performance will realize the students to focus on their impactful courses attentively. Faculty members also guide them to easily understand their course study and find the efficient technique of how they will capable of assisting their good grades on these critical courses. If the instructors and students combinedly focus on their lessons, their CGPA will be higher than the previous technique. Our educational organization should develop a prediction performance system so that, students can easily pass their courses, obtain a good CGPA with their skills, and can join the job sectors to improve the nation's economy. To solve this problem, we develop a generalized framework that helps us predict the student CGPA using both traditional ML and deep learning approaches, only utilizing the student's prior academic results.

Our research aim is to predict student performances by analysis of educational data originating from transcripts that the students observed the proper way to fulfill their expectations. We used the student's transcript data for analysing their previous academic records from a reputed university in Bangladesh. To apply this innovative method, faculties will be capable to understand and improve their learners in academics and restructure course patterns in an organized way.

The prediction of the final academic outcome can help students identify their situations and enhance their current situation accurately. Students can easily understand their laggings through this early prediction system and can get extra guidance from their instructions. It supports the students to overcome their frustration and acquire a good CGPA in their graduation or final result. The institution can focus on backward students and guide them properly by assigning mentors to discuss non-academic perspectives. In this way, faculty members accurately understand the student's patterns and push them to improve their lessons.

The rest of the paper is organized as follows. Section 2 explores previous research works and their outcomes. Section 3 provides the general methods of data mining (such as pre-processing, classification, and prediction) and their evaluation processes. Section 4 determines the result through experimental steps, compares the predicted value with the actual result, and justifies the outcome. Section 5 explains the efficiency of the applied method, the importance of research work, the contribution of this research, and some future works.

2. LITERATURE REVIEW

The process of giving or receiving systematic instructions for acquiring knowledge is called education. Currently, many researchers focus on applying data mining techniques to address issues and make decisions in the field of education. EDM is a developing field connected to advanced methods for exploring unique and increasingly large-scale data. The researchers use different machine learning algorithms to build predictive models. Machine learning is a concept that allows the machine to acquire knowledge and make decisions from prior observations. Machine learning is dissected into two parts such as (a) Supervised and (b) Unsupervised learning. In the new era of the world, EDM has been used in different fields to solve challenges.

Discovery of the knowledge from data sources is known as data mining [7], defined as analysing large information repositories and discovering implicit but adequate latency information. Hidden relationships are related to uncovering data mining and revealing unknown patterns & trends by digging into massive data. As stated by the task performed, the models of data mining are divided into four major categories: (1) association (2) classification (3) clustering, and (4) regression [8-10]. Data mining analysis is the form generally of three techniques: (i) classical statistics (ii) artificial intelligence and (iii) machine learning [11]. Classical statistics are chiefly used for monitoring data, data connections, and vast databases with numeric data for trading [12].

Classical statistics include regression analysis, clump analysis, and distinguish analysis whether artificial intelligence (AI) applies the process to statistical problems. AI uses several techniques such as genetic algorithms, neural computing, weather forecasting etc. [13]. Advanced methods of statistical AI are the combination of advanced statistical methods and AI used for data analysis and knowledge discovery. Machine learning uses several classes of theory such as neural networks [5], symbolic learning, genetic algorithms, and optimization algorithms. Data mining benefits from these technologies but differs from the objective pursued: extracting patterns, describing trends, and predicting behaviour.

Data mining is considered the foundation of scientific research and is used in several environments in the world of scientific research. It is not given enough consideration to examine the environment or its philosophical foundations. This is apparent that the abstract studies of mining data as a scientific investigation field, collection of isolated algorithms patterns alternatively, is helpful for further development of the entire domain. A philosophy layer, a method layer, and an application layer combinedly construct the data mining technique more acceptable. Many types of prime questions focus on mining data above these layers and collectively form

a full epithet of the area. The layered framework is incontestable by applying three sub-fields: classification, measurements, and justify-oriented data mining [14]. Another study defines three popular views of mining data, (a) function-oriented, (b) theory-oriented, and (c) procedure-oriented, respectively [15]. Introduction to the development and aim of data mining, many researchers became concerned about the fundamental issues of mining the data [2], [5], [16]. In [17], the authors mention data management technique is very important to organize any kind of data analysis. In [18], the review literature discusses the security issues of data management and proposed a security model. In [19], the authors mention, data needs pre-process before applying the data mining algorithms. In [20], the authors discuss the data mining technique to find the best possible result from the huge amount of data.

Two decades past, numerous education has been arranged in the concealed pattern in mining data [21]. In this segment, we are going to recapitulate the education discipline shortly. Most of the research in this domain has emphasized utilizing data mining theory. Some sources applied several cases of machine research theory, guessed the ability of the freshers, chose the actual catalogue for the future idea, and discovered the discussed few of the recent and pertinent studies. Using only the results from the prior subjects, decision trees, a well-known machine learning classifier method, are used to estimate student cumulative GPA [16]. They collected the dataset including student final GPA and CGPA in all programs from the student transcript, but their focus was only on the mandatory courses. They discretized the numerical utility of the student grade into categorical groups that indicated a high level of grade inflation. They have also found a relation between final GPA and the progress of students in their future [16]. Sushmita has proposed a model [22] for predicting student outcomes to inspect and classify student details based on academic records, and the K-mean clustering algorithm has been used. This analysis considered the most accurate attributes from the total 24 attributes where CGPA, attendance, GPA of different semesters, education medium, and board types have been selected. In another study [22], authors predict the student ability by comparative analysis of efficient classification algorithms, namely naive bayes, multi-layer perceptron, SMO, decision tree, and REP tree. They obtained a high influence using a decision tree with an accuracy of 97%.

The assessment data are applied to data mining algorithms to determine the success rate in a program (either passed or failed). The acting of the learning theory is evaluated on its predictive quality, ease of study, and user-friendly characteristics. After finishing the data collection and data pre-processing, the data mining theory is applied. This methodology is used to help students of gaining advanced techniques from proceedings, of the international conference on intelligent tutoring systems, and the Educational Data Mining Workshop (EDMW).

3. METHODOLOGY

In our novel research, we have followed an appropriate methodology for conducting our experiment. Multiple workflows for classification, regression, and pattern analysis have been preserved. We suggested a generalized framework for forecasting each student's class level and pattern analysis in the first phase. We also developed a DNN architecture for predicting the final CGPA of undergraduate students.

At first, we collected data from survey responses. After collecting the data, we pre-processed the dataset by applying manual techniques. We applied six classification algorithms to our pre-processed dataset and evaluate the algorithms with some pre-defined evaluation metrics. Again we applied a DNN architecture in our dataset and used some evaluation metrics to assess the outcome. Finally, we observed the prediction results based on the evaluation matrices. In figure 1 our proposed research methodology is described, and the steps of our research work are mentioned clearly.

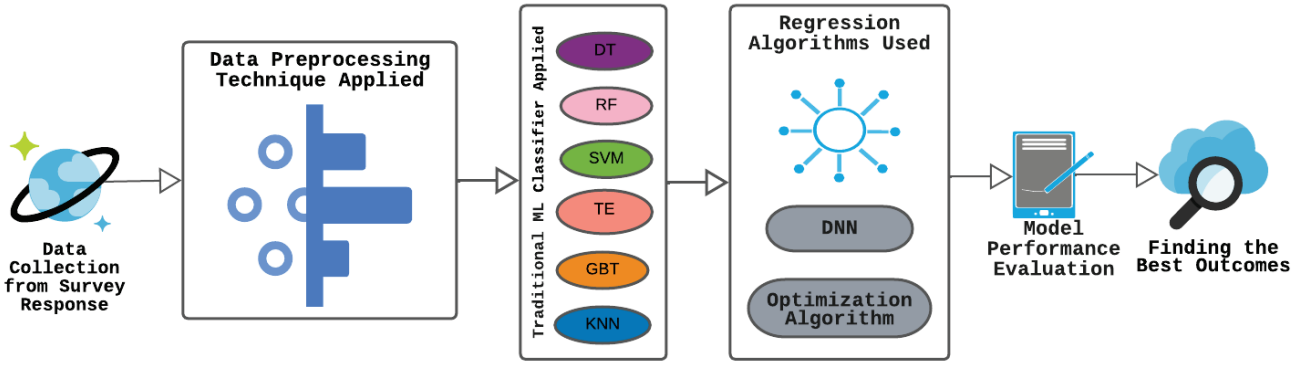

Fig. 1. Proposed research methodology

We can acquire a good result from the model that has been applied, as directly impacts the model's ability to learn. Though, to tackle a real-world problem with the help of machine learning, it is critical to pre-process the raw data accurately. For data inconsistency, most of the time real data can't access easily even if it is critical for pre-processing. Figure 2 shows the pre-processing of the data where procedures are briefly addressed. We have manually handled the missing data and all particular instances labelled by column. We have changed the data using the dataset versioning systems. After transforming the data we extracted the data using the feature engineering method. Then move into the feature encoding concept to scale the data.

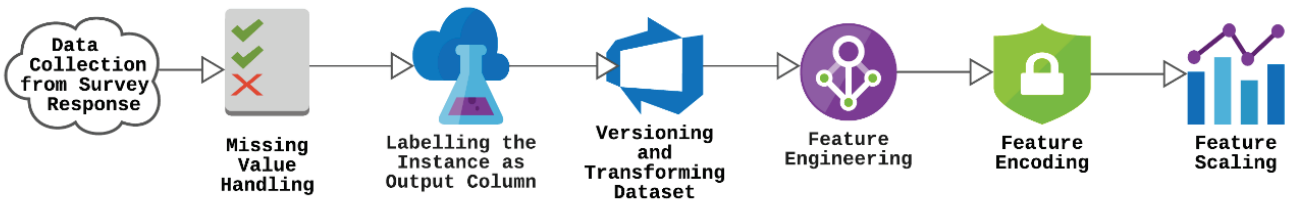

Fig. 2. Pre-processing of the dataset

Actual data frequently has some missing values, which are generally cleaned during the experimental data pre-processing. Before dealing with the missing data, we must first learn about its nature of the missing data. Data might be absent in various ways. We find missing values when collecting the data with empty entry, malfunction of data collection, or spoiling of data items which has been collected before. Figure 3 shows the different types of missing data.

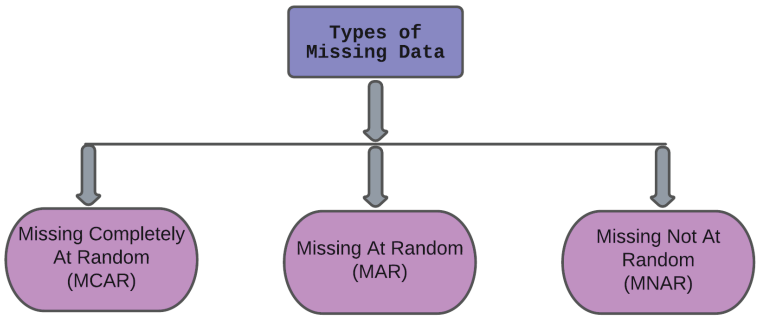

Fig. 3. Types of missing data

In the first two types (MCAR & MAR), that is safe to delete data with missing values based on their frequency, however in the third instance (MNAR), eliminating observations with missing values can determine to lead forward based on the concept of biasing. We have collected a total of 398 instances including some missing data of MCAR type. We have removed the missing values from our dataset and getting 372 instances left for the next processing. Finally, our prediction model identified the difference between actual grade and predicted grade from the pre-processed dataset using both traditional machine learning classifiers and deep neural networks. The visual architecture of our proposed DNN model is depicted in figure 4.

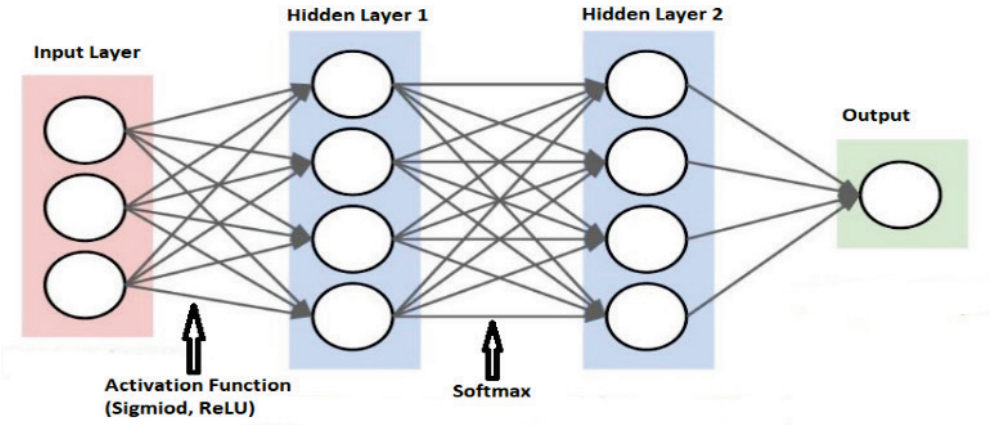

Fig. 4. DNN Architecture

At first, hyperparameters are configured to design a neural-network architecture. Hyperparameters are the number of neurons that are in input and output layers. The training dataset determines the value of these hyperparameters. In this research, the structure of our training dataset is (261, 43), indicating that there are 261 instances and 43 attributes. From this structure of the training dataset, 43 input features have been specified by 43 neurons in the input layer.

Determination of the number of hidden layers and the number of neurons in each hidden layer is essential to construct the architecture of a neural network. According to the research by [23], an approximation of any function can be made by networks with a single hidden layer. With rational activation functions, networks with two hidden layers may characterize a despotic boundary of the decision to despotic accuracy. Furthermore, networks with more than three hidden layers can learn complex designs and deliver automated

feature engineering. If the number of hidden layers is more than two, the model turns to complex architecture. Sometimes, a neural network with more hidden layers makes the model computationally inefficient. As a result, we created a neural network architecture with two hidden layers.

A rule of thumb has been used [24] to find out the number of neurons in the hidden layer. In this research, 64 neurons were used for the first hidden layer since our 43 is the dimension of our final input's first hidden layer. We used 32 neurons in the second hidden layer. We have used the ReLU activation function which is one of the essential things of the neural network.

$$f(x) = \max\ (0, x) \quad (1)$$

Where x is the input of each neuron. Activation function map the data between input and response data. We used six different evaluation indicators for traditional machine learning classifiers for assessing our predictive performance. Three different assessment metrics in deep learning algorithms evaluate the effectiveness of our model based on error measurement. All of the indicators were chosen based on their use and applicability.

4. RESULTS AND DISCUSSIONS

We prepare our dataset from an online survey collected from student's transcript data acquired in the department of marketing of a prestigious institution. We gather information from students who have already graduated or almost complete all the courses. Our dataset has a total of 398 instances. Each instance contains the information students name, id, gender, all the subjects information with the grade, GPA of each semester, the status based on GPA, and overall CGPA. After data collection, we eliminate useless attributes like student name, student id, course descriptions, etc., to minimize the computation cost. Then, pre-processing steps have been applied to the dataset to handle the missing value of that dataset.

The dataset was evaluated and compared with six different classification methods, including GBT, RF, TE, DT, SVM, and KNN, to indicate the performance class. Our pre-processed dataset is used in different classification methods to predict the student's performance. Precision, recall, specificity, and the f1-measure are used to determine the particular method's results. We have found different classification performances for an individual class. Finally, we evaluated the accuracy and cohen's kappa with an overall high degree of accuracy.

Table 1 represents the overall evaluation with a high degree of accuracy. In the table, random forest algorithms find out the highest accuracy of student performance with the record-level f-measure and sensitivity.

Table 1. Different evaluation models with a high degree of accuracy

Algorithms Applied	Class	Precision	Sensitivity	Specificity	F-measure	Accuracy	Cohen's Kappa
Decision Tree (DT)	*Honors*	0.300	0.300	0.981	0.300	0.884	0.361
	First Class	0.935	0.937	0.405	0.936		
	Second Class	0.462	0.444	0.959	0.453		
Random Forest (RF) *	*Honors*	0.833	0.500	0.997	0.625	**0.941 ***	**0.598 ***
	First Class	0.946	0.991	0.486	0.968		
	Second Class	0.867	0.481	0.994	0.619		
Support Vector Machine (SVM)	*Honors*	1.000	0.500	1.000	0.667	0.914	0.491
	First Class	0.944	0.961	0.486	0.953		
	Second Class	0.500	0.481	0.962	0.491		
Tree Ensemble (TE)	*Honors*	1.000	0.300	1.000	0.462	0.938	0.544
	First Class	0.938	0.997	0.405	0.967		
	Second Class	0.923	0.444	0.997	0.600		
Gradient Boosted Tree (GBT)	*Honors*	0.556	0.500	0.989	0.526	0.927	0.515
	First Class	0.940	0.982	0.432	0.961		
	Second Class	0.846	0.407	0.994	0.550		
K-Nearest Neighbors (KNN)	*Honors*	0.778	0.700	0.994	0.737	0.919	0.435
	First Class	0.932	0.982	0.351	0.956		
	Second Class	0.600	0.222	0.988	0.324		

Figure 5 shows the overall performance of different evaluation metrics. In this figure, we calculate every evaluation metric individually for our applied different classification methods. If we notice carefully, random forest performs the highest results where the decision tree generates the lowest score than the other applied methods.

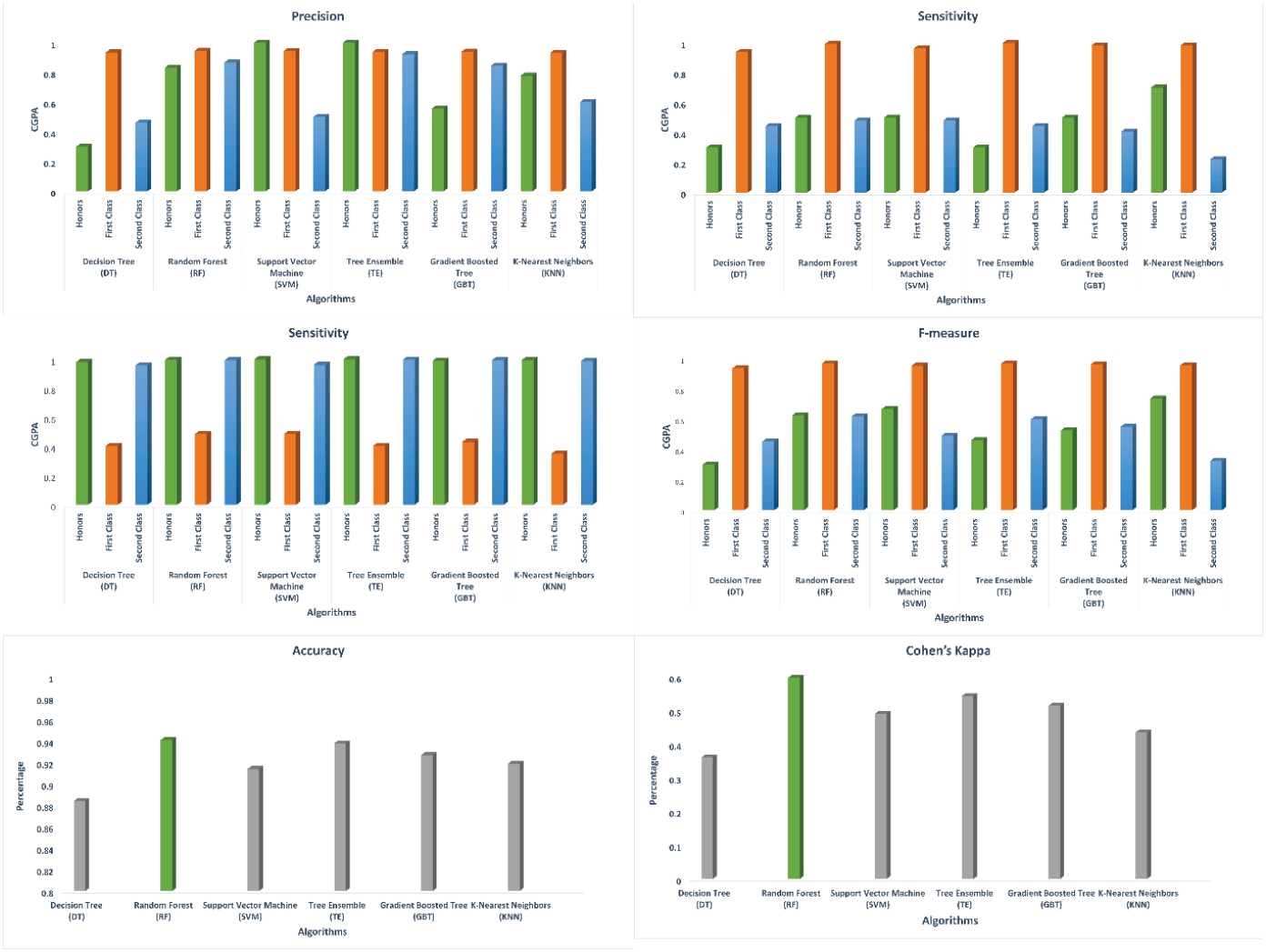

Fig. 5. Performance plot of different classifiers with evaluation metrics

After completing the training session, we applied our proposed method DNN to make predictions on our dataset. Table 2 represents the difference between the actual and predicted CGPA. Our model performs some CGPA predictions on test data after the training phase of DNN.

Table 2. Prediction results of the proposed model

Gender	Actual CGPA	Predicted CGPA	Difference
M	3.52	3.49	0.03
F	3.54	3.52	0.02
M	3.44	3.46	0.02
F	3.51	3.51	**0.00 ***
M	3.37	3.33	0.04
F	3.75	3.77	0.02
M	3.36	3.28	0.08
F	3.71	3.62	0.09
M	3.44	3.36	0.08
F	3.72	3.68	0.04

In this research, we used three error metrics such as MSE, MAE, and MAPE to evaluate the performance of the proposed neural network model. We used the MAE as the loss function and learned our model for 50 epochs. Figure 6 represents the train vs. validation accuracy graph based on the dataset and shows the model's learning rates through the epochs. As can be displayed, the two-axis have converged at the 10th epochs between loss vs. epochs. The converging line indicates that the two-axis line has been quite compatible in converging between loss and epochs, indicating that the model is well-fitted.

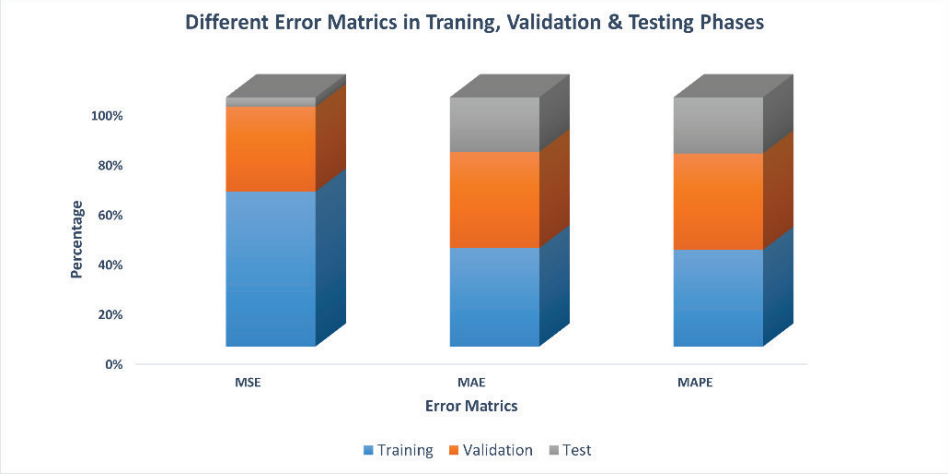

Fig. 6. Error ratio of the Training, Testing and Validation data

In the training and validation phases, the model is used to predict the final CGPA from the test dataset. Our pre-processed dataset consists of a total of 372 instances divided into two major groups: 327 instances in the training dataset and 45 instances in the test dataset. Our model's training, validation, and test results are shown in Table 3. The training and validation outcome values are close to each instance. It indicates that our applying model has learned properly without overfitting from the training dataset. Surprisingly, the outcomes are relatively consistent across all phases, with slight fluctuation.

Table 3. Performance of the proposed model

	Mean Squared Error (MSE)	Mean Absolute Error (MAE)	Mean Absolute Percentage Error (MAPE)
Training	0.129	0.118	3.525
Validation	0.069	0.114	3.505
Test	0.007	0.064	1.951

We plot 10 instances from our test dataset into a line graph for a better understanding of the variances in actual and predicted CGPA. Figure 7 shows the differences between predicted and actual CGPA values. The model fits perfectly and accurately predicts the actual value for most samples. The pattern analysis and error metrics allow us to draw the conclusion that the applied model can accurately predict the overall CGPA of undergraduate students.

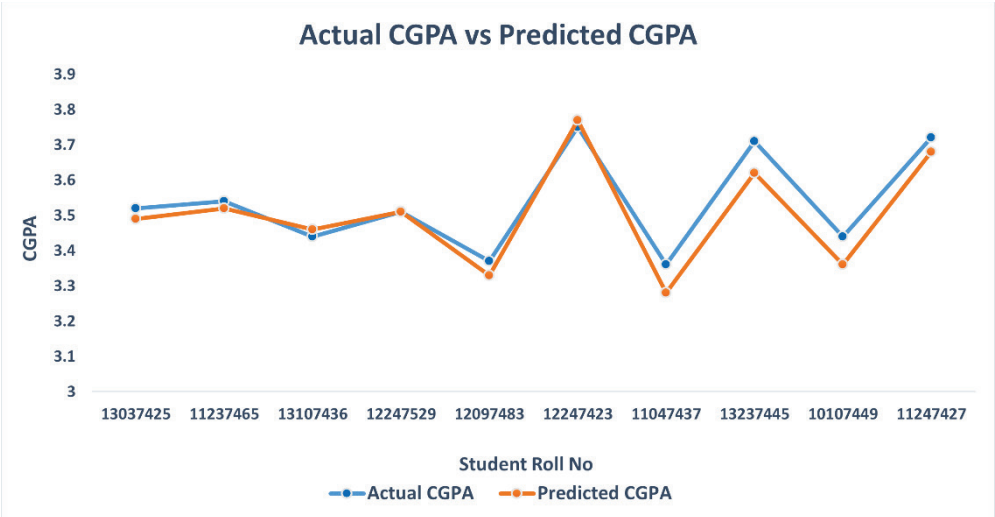

Fig. 7. Actual CGPA vs Predicted CGPA

The experimental phase evaluates the classification and regression algorithms to determine the highest result by predicting student performance. The model analyses the actual and predicted values properly from different category students. Random forest outperforms other models with 94.1% accuracy in all these algorithms. To maintain this pattern of algorithms and evaluate the method, undergraduate students can get proper guidance from their instructors during their graduation period.

5. CONCLUSION

The goal of this novel research is to perform prediction and pattern recognition to help undergraduate students to enhance their results. First, a generalized framework is applied that can assist faculties or teachers in identifying and correctly guiding students who are at risk. To classify the student performance and generate the outcome depending on their academic history in each semester, GBT, RF, TE, DT, SVM, and KNN algorithms were used. A DNN model has been introduced to predict the final CGPA of undergraduate students with a lower error ratio than standard models. Three evaluation measures, MSE, MAE, and MAPE, were used to assess the performance of the DNN model. MAE is utilized as the loss metric to look at the training and validation losses. This study has promising results, such as 0.007 for MSE, 0.064 for MAE, and 1.951 for MAPE. Furthermore, compared to the baseline decision tree model, the approach of this study reduced MSE, MAE, and MAPE errors by 0.0144, 0.0429, and 6.039, respectively. The majority of the predicted points are extremely near to the actual values.

The performance of the proposed model has demonstrated compatible learning from the transcript data and promising prognostic performance. Additional factors like environment, study behaviour, study schedule, family assistance, and disqualification need to be incorporated to improve the output. In the area of NLP, BERT is a well-known deep learning technique with high algorithmic efficiency. However, a high configuration device setup and significant computation cost are required. Future studies will use the BERT model to forecast students' performance in order to improve the performance of the proposed strategy.

REFERENCES

[1] Peña-Ayala, A. (2014). Educational data mining: A survey and a data mining-based analysis of recent works. Expert Systems with Applications, 41(4, Part 1), 1432-1462. Retrieved from https://www.sciencedirect.com/science/article/pii/S0957417413006635 doi: 10.1016/j.eswa.2013.08.042

[2] Hand, D. J., Mannila, H., & Smyth, P. (2007). Principles of data mining. Drug Safety, 30, 621-622. Retrieved from https://pubmed.ncbi.nlm.nih.gov/17604416 doi: 10.2165/00002018-200730070-00010

[3] Anjir Ahmed Chowdhury, S. K. S. M. R., Argho Das, & Hasan, K. T. (2021, May). Sentiment analysis of covid-19 vaccination from survey responses in bangladesh. Cognitive Computation. Retrieved from https://www.researchsquare.com/article/rs-482293/v1 doi: 10.21203/rs.3.rs-482293/v1

[4] Sukhija, K., Jindal, M., & Aggarwal, N. (2016). Educational data mining towards knowledge engineering: a review state. International Journal of Management in Education, 10(1), 65-76. Retrieved from https://www.inderscienceonline.com/doi/abs/10.1504/IJMIE.2016.073362n doi: 10.1504/IJMIE.2016.073362

[5] Islam Rifat, M. R., Al Imran, A., & Badrudduza, A. S. M. (2019). Edunet: A deep neural network approach for predicting cgpa of undergraduate students, 1-6. Retrieved from https://ieeexplore.ieee.org/document/8934616 doi: 10.1109/ICASERT.2019.8934616

[6] Asian Development Bank (ADB). (2011). Higher education across Asia: An overview of issues and strategies. Manila: ADB.

[7] Al-Barrak, M. A., & Al-Razgan, M. (2016, July). Predicting students final gpa using decision trees: A case study. International Journal of Information and Education Technology, 6(7), 528-533. Retrieved from http://www.ijiet.org/show-74-837-1.html doi: 10.7763/IJIET.2016.V6.745

[8] Hui, S., & Jha, G. (2000). Data mining for customer service support. Information Management, 38(1), 1-13. Retrieved from https://www.sciencedirect.com/science/article/pii/S037872060000051 doi: 10.1016/S0378-7206(00)00051.

[9] Kao, S.-C., Chang, H.-C., & Lin, C.-H. (2003). Decision support for the academic library acquisition budget allocation via circulation database mining. Information Processing Management, 39(1), 133-147. Retrieved from https://www.sciencedirect.com/science/article/abs/pii/S0306457302000195 doi: 10.1016/S0306-4573(02)00019-5

[10] Nicholson, S. (2006). The basis for bibliomining: Frameworks for bringing together usage-based data mining and bibliometrics through data warehousing in digital library services. Information Processing Management, 42(3), 785-804. Retrieved from https://www.sciencedirect.com/science/article/pii/S0306457305000658 doi: 10.1016/j.ipm.2005.05.008

[11] Girija, N., & Srivatsa, S. (2006). A research study: Using data mining in knowledge base business strategies. Information Technology Journal, 5(3), 590-600. Retrieved from https://scialert.net/abstract/?doi=itj.2006.590.600 doi: 10.3923/itj.2006.590.600

[12] Hand, D. J. (1998). Data mining: Statistics and more? The American Statistician, 52(2), 112-118. Retrieved from https://www.tandfonline.com/doi/abs/10.1080/00031305.1998.10480549 doi: 10.1080/00031305.1998.10480549

[13] Iyanda, A., Ninan, D., Ajayi, A., & Anyabolu, O. (2018, jun). Predicting student academic performance in computer science courses: A comparison of neural network models. International Journal of Modern Education and Computer Science, 10, 1-9. Retrieved from https://www.mecs-press.org/ijmecs/ijmecs-v10-n6/v10n6-1.html doi: 10.5815/ijmecs.2018.06.01

[14] Daud, A., Aljohani, N. R., Abbasi, R. A., Lytras, M. D., Abbas, F., & Alowibdi, J. S. (2017). Predicting student performance using advanced learning analytics., 415–421. Retrieved from https://doi.org/10.1145/3041021.3054164 xdoi: 10.1145/3041021.3054164

[15] Dietterich, T. G. (2000, jan). Ensemble methods in machine learning. Multiple Classifier Systems: First International Workshop, MCS 2000, Lecture Notes in Computer Science, 1-15. Retrieved from https://link.springer.com/chapter/10.1007%2F3-540-45014-9_1 doi: https://doi.org/10.1007/3-540-45014-91

[16] Mueen, A., Zafar, B., &Manzoor,U. (2016, 11).Modeling and predicting students' academic performance using data mining techniques. International Journal of Modern Education and Computer Science, 11, 36-42. Retrieved from https://www.mecs-press.org/ijmecs/ijmecs-v8-n11/IJMECS-V8-N11-5.pdf doi: 10.5815/ijmecs.2016.11.05

[17] Bhuiyan, M. N., Masum Billah, M., Saha, D., Rahman, M., Kaosar, M., et al. (2022). Iot based health monitoring system and its challenges and opportunities. In Ai and iot for sustainable development in emerging countries (pp. 403–415). Springer.

[18] Bhuiyan, M. N., Rahman, M. M., Billah, M. M., & Saha, D. (2021). Internet of things (iot): A review of its enabling technologies in healthcare applications, standards protocols, security and market opportunities. IEEE Internet of Things Journal.

[19] Billah, M. M., Bhuiyan, M. N., & Akterujjaman, M. (2021). Unsupervised method of clustering and labeling of the online product based on reviews. International Journal of Modeling, Simulation, and Scientific Computing, 12(02), 2150017.

[20] Hossain, M. A., Ali, M. A., Kibria, M. G., & Bhuiyan, M. N. (2013). A survey of e-commerce of bangladesh. International Journal of Science and Research, 2(2), 150–158.

[21] Chaudhury, P., & Tripathy, H. (2017, jan). An empirical study on attribute selection of student performance prediction model. International Journal of Learning Technology, 12(3), 241–252. Retrieved from https://doi.org/10.1504/IJLT.2017.088407 doi: 10.1504/IJLT.2017.088407

[22] Sumitha, R., & Vinothkumar, E. S. (2016, jun). Prediction of students outcome using data mining techniques. International Journal of Scientific Engineering and Applied Science (IJSEAS), 2(6).

[23] Heaton, J. (2008). Introduction to neural networks for java, 2nd edition.

[24] Moolayil, J. (2018). Learn keras for deep neural networks: A fast-track approach to modern deep learning with python. Retrieved from https://www.oreilly.com/library/view/learn-keras-for/9781484242407/