Predictive Modeling Using Artificial Intelligence Algorithms to Forecast Academic Grades

" النمذجة التنبؤية باستخدام خوارزميات الذكاء الاصطناعي للتنبؤ بالدرجات الأكاديمية "

Nora Elrashidy[1], Marwa El-Sedik[1], Zeinab Hassan[1], Abeer Saber[2], Esraa Hassan[1]

[1]*Faculty of Artificial Intelligence, Kafrelsheikh University, Kafrelsheikh 33511, Egypt.*

[2]*Faculty of Computers and Artificial Intelligence, Damietta University, Damietta 34517, Egypt.*

Volume Three - Issue Seven - February 2025

المجلد الثالث - العدد السابع – فبراير ٢٠٢٥

ISSN-Online: 2812-6122 ISSN-Print: 2812-6114

موقع المجلة على بنك المعرفة المصري

https://aiis.journals.ekb.eg/contacts?lang=ar

Abstract: Accurate academic grade forecasting is crucial for improving student outcomes and optimizing resource allocation in educational settings, but traditional methods often lack complexity. In this paper, we explore using various algorithms for predictive modeling to forecast academic grades. The primary goal is to identify the most effective algorithm for predicting student performance, which can assist educators and policymakers in making informed decisions. Algorithms such as Linear Regression (LR), Decision Trees (DT), Random Forest (RF), and Gradient Boosting Regression (GBR) were utilized to build the models. Each was assessed using metrics including training and testing scores, Mean Squared Error (MSE), Mean Absolute Error (MAE), Median Squared Error (MedSE), and R-squared (R^2).The findings revealed that Gradient Boosting Regression (GBR) achieved the highest accuracy, with a training score of 0.999, a testing score of 0.96, an MSE of 0.5068, an MAE of 0.258, and an R^2 of 0.995. These results indicate that GBR surpasses the other models in predicting academic performance, offering a reliable tool for grade forecasting and supporting educational planning and intervention strategies.

Keywords: Academic Grade Forecasting, Predictive Modeling, Educational Planning

المستخلص : تستعرض هذه الورقة خوارزميات مختلفة للنمذجة التنبؤية للتنبؤ بالدرجات الأكاديمية. وكان الهدف الرئيسي هو تحديد الخوارزمية الأكثر فعالية للتنبؤ بأداء الطلاب، مما يمكن أن يساعد المعلمين وصانعي السياسات والغايات (DT) ، وأشجار القرار(LR) في اتخاذ قرارات مستنيرة. تم استخدام خوارزميات مثل الانحدار الخطي لبناء النماذج. تم تقييم كل منها باستخدام معايير (GBR) ، والانحدار باستخدام التعزيز التدريجي(RF) العشوائية ، متوسط الخطأ (MAE) ، متوسط الخطأ المطلق(MSE) تشمل درجات التدريب والاختبار، متوسط الخطأ المربع (GBR) كشفت النتائج أن الانحدار باستخدام التعزيز التدريجي .(R^2) ، ومعامل التحديد(MedSE) المربع الوسيط بمقدار MAE ،٠,٥٠٦٨ حققMSE أعلى دقة، حيث كانت درجة التدريب ٠,٩٩٩، درجة الاختبار ٠,٩٦ ، يتفوق على النماذج الأخرى في التنبؤ GBR بمقدار ٠,٩٩٥. تشير هذه النتائج إلى أن R^2 بمقدار ٠,٢٥٨، و بالأداء الأكاديمي، مما يوفر أداة موثوقة للتنبؤ بالدرجات ودعم استراتيجيات التخطيط والتدخل التعليمي .

الكلمات المفتاحية : التنبؤ بالدرجات الأكاديمية، النمذجة التنبؤية، التخطيط التربوي

1.Introduction

Forecasting academic grades using machine learning algorithms represents a transformative approach in education, offering predictive insights that can revolutionize student support and educational outcomes. Machine learning algorithms may predict future grades with amazing accuracy by examining a wide range of characteristics, including past academic achievement, attendance records, study habits, and even social elements. These algorithms can uncover subtle patterns and connections in student data, allowing educators to identify failing children early and apply targeted interventions to improve their academic performance. Through the proactive application of machine learning, educational institutions can move beyond reactive tactics and adopt a preventative approach that promotes student progress and achievement [1-2]. Machine learning algorithms predict academic grades and provide a better understanding of the underlying elements that influence student performance. These models can unearth hidden insights that help educators customize their teaching approaches to effectively meet the needs of individual students by assessing a varied variety of data points such as student demographics, learning styles, and extracurricular activities. Furthermore, machine learning may help educators create individualized learning paths that guide students in receiving tailored support based on their individual strengths, limitations, and learning preferences. This data-driven strategy enables educators to make educated decisions that improve student engagement, motivation, and academic performance [3-4]. Integrating machine learning in academic grade forecasting can revolutionize education by fostering a culture of data-driven decision-making and personalized student support. These algorithms predict future educational outcomes and enable educators to optimize resource allocation, refine curriculum design, and create targeted interventions that address specific student needs [21-23]. By harnessing the power of machine learning, educational institutions can move towards a more proactive, student-centric approach to education, ensuring that each student receives the personalized attention and support necessary to thrive academically and reach their full potential [5]. The main objective of this paper is to develop the most effective algorithm for forecasting student achievement, which will help educators and policymakers make educated decisions. The models were built using LR, DT, RF, and GBR algorithms. Each was evaluated using training and testing scores, MSE, MAE, MedSE, and R^2 measures.GBR had the highest accuracy, with a training score of 0.999, a testing score of 0.96, an MSE of 0.5068, an MAE of 0.258, and an R^2 of 0.995. These findings show that GBR outperforms other models in predicting academic achievement, providing a dependable tool for grade predictions and aiding educational planning and intervention measures.

The rest of this paper is organized as follows: Section 2 discusses related works and recent literature review. Section 3 provides the proposed framework for predicting student performance using artificial intelligence methods. Section 4 explores results and experiments. Section 5 provides a paper conclusion.

2.Related Works

In this section, we review the key related works relevant to literature which aims to apply machine learning methods in educational research to address existing gaps in the literature. This approach holds great significance for educational studies, allowing researchers to leverage high-dimensional data to tackle issues beyond the capabilities of traditional statistical models and enhance prediction accuracy when dealing with complex datasets. Several relevant works suggest models, as illustrated in Table 1. H. Bai et al. [1] utilized machine learning with multiple procedures, incorporating the bootstrap resampling technique in random forest functions. This method was employed to assess the importance of predictors and determine the significance of variables in predicting students' final math grades (G3). The use of random forests provided robustness against overfitting and offered insights into variable importance, although it may be computationally intensive and less interpretable than simpler models. M. S. Croock et al. [2] proposed a method for preprocessing Educational Data Mining datasets and selecting features using a hybrid approach that combines filter and wrapper techniques. For filter-based feature selection, they used statistical analysis methods such as Pearson correlation and information gain. The wrapper method employed a neural network. This hybrid approach balances the computational efficiency of filter methods with the accuracy of wrapper methods, though it can be more complex to implement and interpret. A. Goodarzi et al. [3]developed two predictive models that identified family and study time as the most influential factors on students' academic success in mathematics and Portuguese language classes. These models provided actionable insights for educational interventions, highlighting key predictors of academic performance. However, the focus on specific subjects may limit the generalizability of the findings to other areas. S. T. Ahmed et al. [4] made a significant contribution by developing a new multi-objective decision tree called the Decisive Decision Tree (DDT), designed for feature selection and classification. This model achieved a high accuracy of 92%, demonstrating the effectiveness of multi-objective optimization in educational data mining. The DDT method enhances interpretability and accuracy, though the complexity of multi-objective optimization might pose challenges in implementation. Cem Özkurt [5]utilized interpretable artificial intelligence models, such as InterpretML, to analyze datasets. The study evaluated the results to understand the factors influencing student success. The use of interpretable AI models provides transparency and insights into the decision-making process, making it easier for educators to apply findings in practical settings. However, these models might sacrifice some predictive power for the sake of interpretability. These studies illustrate the diverse applications and methodologies of machine learning in educational research, each with its advantages and disadvantages in terms of accuracy, interpretability, and computational demands.

Table 1: The state-of-the-art of related study work.

Authors / Year	**Method**	**Evaluation Measures (%)**	**Limitation**

H. Bai et al.[1] /2022	Random Forst	R2 =.0576 MAE2 =3.623	Used one type of machine learning model for analyzing educational data
M. S. Croock et al.[2] /2019	Neural network	ACC= 0.92	Used all featured instead of selected a minimal subset of features with high predictive power.
A. Goodarzi et al.[3] / 2024	Random Forest	R2=0.42 MAE2 =5.36	The dataset used is not extensive enough to improve predictions or identify unique influences on student performance across various academic subjects
S. T. Ahmed et al. [4] /2020	Decisive Decision Tree	Acc =0.92	research and testing are needed to assess its scalability and effectiveness across diverse and more complex datasets.
Cem Özkurt [5]/2024	InterpretML	Acc=0.742 Recall =0.888	non-linear relationships between variables as effectively. Reduced predictive accuracy.

3.Proposed Work

In this section, we propose a predictive model to forecast academic grades, as shown in Figure 1, utilizing various algorithms: linear regression (LR), decision trees (DT), random forest (RF), and gradient boosting regression (GBR). Each algorithm offers distinct advantages in handling diverse data and relationships. The proposed methodology encompasses several key stages:

3.1. Data Collection and Preprocessing

Data collection and preprocessing are crucial in developing a predictive model to forecast academic grades. These steps ensure the data is clean, relevant, and suitable for analysis, which is vital for building accurate and reliable models. The following subsections detail the main activities involved in these processes that are illustrated in Table 1.

Table 2: The general description for data collection and preprocessing.

Data Processing	**Description**
Data Gathering	Acquire a comprehensive dataset containing students' academic records, including previous grades, attendance, participation, demographic information, and other relevant attributes.
Data Cleaning	Address missing values, eliminate duplicates, and rectify any inconsistencies.
Feature Engineering	Develop new features from existing data to enhance predictive capability. This may involve creating interaction terms, polynomial features, and domain-specific metrics.
Data Splitting	Divide the dataset into training, validation, and test sets to ensure unbiased evaluation.

The following algorithm contributes uniquely to predicting academic performance, with Random Forest and Gradient Boosting Regression offering advanced methods for handling complex datasets and improving forecast precision that illustrate in Table2.

Table 3: The used algorithms for predictive modeling using artificial intelligence algorithms to forecast academic grades.

Algorithms	**Description**
Linear Regression	Establishing a baseline performance using linear regression to identify linear relationships between input features and grades.
Decision Trees	Implementation of a decision tree to capture non-linear relationships and feature interactions.
Random Forest	Constructing a random forest model to improve prediction accuracy by averaging multiple decision trees,

	thus reducing overfitting.
Gradient Boosting Regression	Applying gradient boosting to iteratively enhance performance by combining weak learners to minimize prediction errors.

Training each model using the training dataset, employing techniques like cross-validation to prevent overfitting with performing hyperparameter tuning using grid search or random search to identify optimal settings that enhance model performance [8-13].

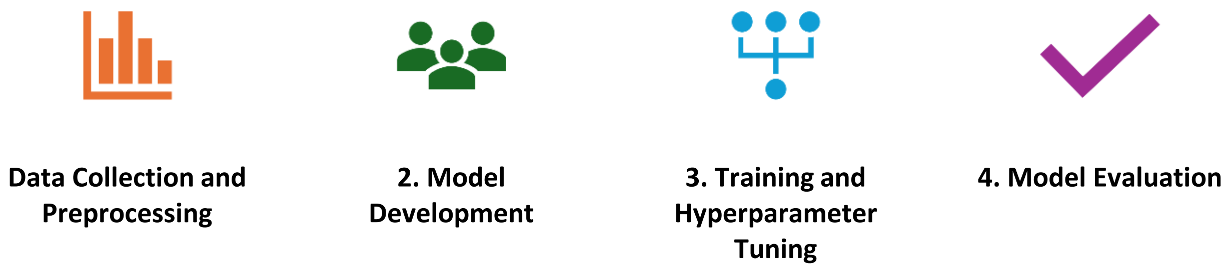

Figure 1: Key Stages of the Proposed Methodology

4.Results and Experiments

4.1 Data set description

The applied dataset describes the final grades students obtained throughout their school year in the math course. G3 is the label or in this case the final grade while the other columns are the features or the inputs.
The feature for the applied dataset is described in Table 3.

Table 3. The applied dataset features.

Feature Name	Feature Type	Feature description
grade_status	Boolean	The status of the student's school grade shows whether the student passed or failed.
sex	Boolean	The gender of the student
age	Numeric	The age of the student.
Medu	Numeric (1-5)	The level of education of the student's mother.

Fedu	Numeric (1-5)	The level of education of the student's father
travel time	Numeric	The duration of time the student spends traveling from home to school.
Study time	Numeric	The amount of time the student spends studying per week.
failures	Numeric	The count of failures the student has experienced.
schoolsup	Boolean	This Boolean feature indicates whether the student receives additional support from the school.
goout	Numeric	The frequency of the student's outings with friends.
Dalc	Numeric	The student's alcohol consumption on weekdays.
Walc	Numeric	The student's alcohol consumption on weekends.
health	Numeric	The student's perceived health status
absences	Numeric	The number of school absences the student has had.
internet_availability	Boolean	This Boolean feature indicates whether the student has internet access at home.
b_Pstatus	Boolean	The represents the cohabitation status of the student's parents, indicating whether they live together or are divorced.
paid_classes	Boolean	This feature indicates whether the student takes additional paid classes
educational_support,	Boolean	This feature indicates whether the student receives additional support from their family.
G3	Numeric	This target feature represents the total grade achieved by the student.

4.2. Experimental analysis

In this section, the performance for the applied algorithms is presented. For more accurate results, the data is first pre-processed [14-16]. The histogram after pre-processing is shown in Figure 2. The distribution plots in Figure 3 show the number of students passed in G1, G2, and G3 while the distribution plot with normal distribution is presented in Figure 4. For improving the education method in the schools, the percentage of students who failed and passed is estimated as shown as in Figure 5.

Figure 6 shows a boxplot that describes all the features in the dataset. Figure 7 depicts the heatmap correlation matrix, which identifies the most significant determinants of grade predictions. Several tests were undertaken to determine the performance of the proposed model on the dataset. The applied dataset was divided to 70% and 30% for the training and testing tasks, respectively [17-20]. From Table 3, it can be observed that the GBR achieved the best training score with 0.999, while the DT is the second best. On the other hand, the DT algorithm achieves the best testing score results with 0.991, while the LR achieves the second best. The results presented in Table 3 confirm that the LR algorithm achieves N/A, 5.643, 1.0305, and 0.8860 for MSE, MAE, MedSE, and R2 Score, respectively. Also, the DT algorithm achieves 1.059, 0.314, 0.0001, and 0.990 for the same metrics, while the GBR algorithm obtains 0.5068, 0.258, .00824, and 0.995, respectively. The results of the investigation show that the suggested model outperforms the DT algorithm in terms of testing scores.

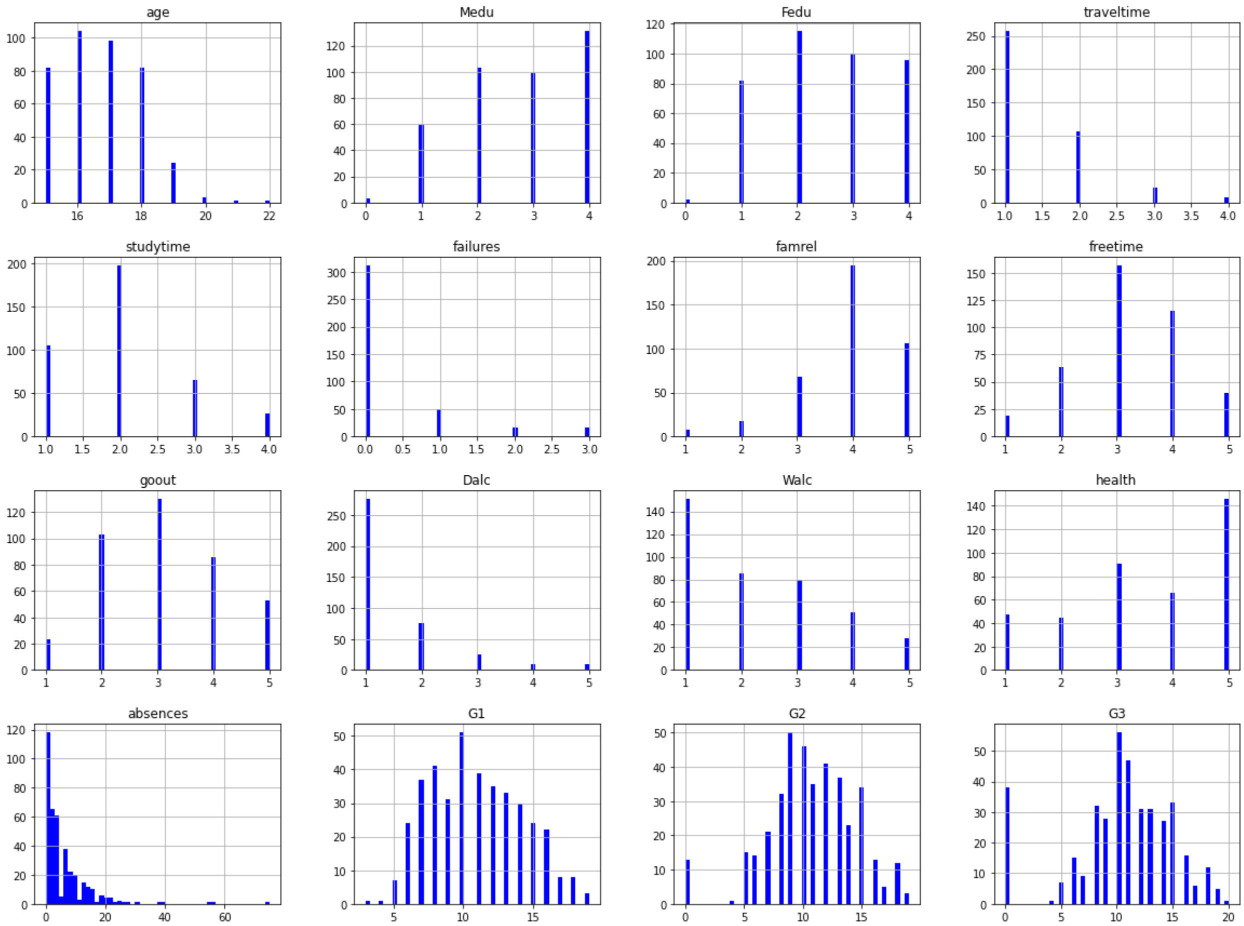

Figure 2: The histogram for the pre-processed data

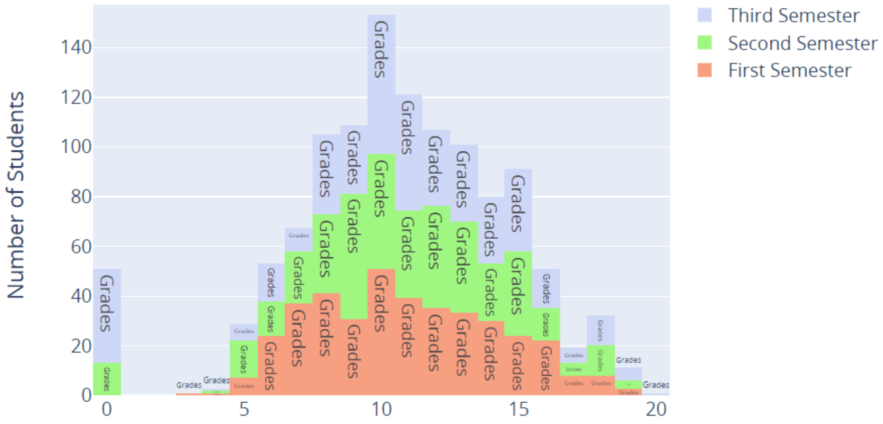

Figure 3: The student grades distribution plot

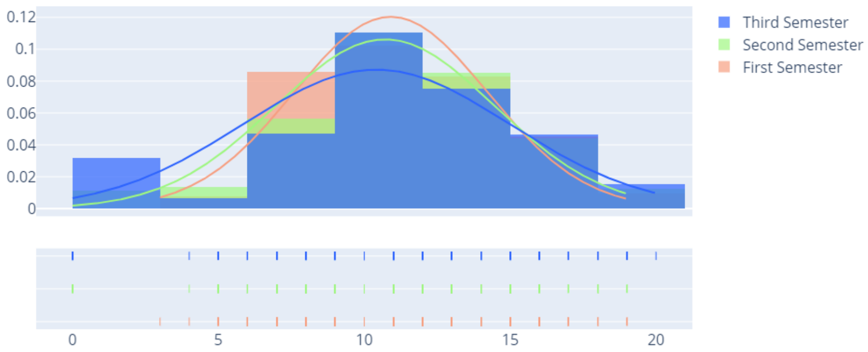

Figure 4. distribution plot with normal distribution

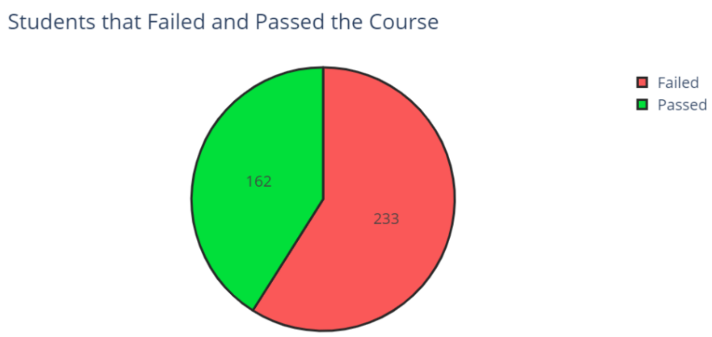

Figure 5. the percentage of students failed and passed

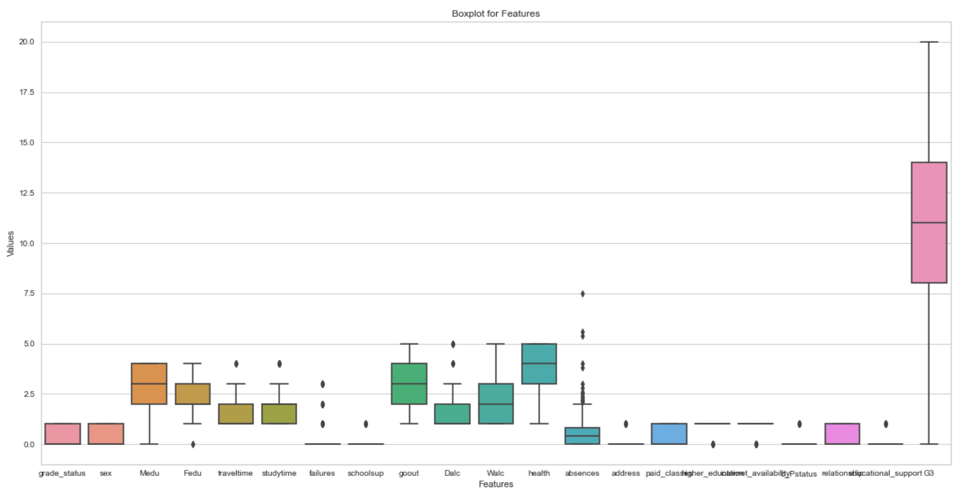

Figure 6. The Boxplot for features

The superior performance of the GBR model can be attributed to its ability to handle complex data structures and interactions between variables. Unlike traditional methods such as Linear Regression, which assumes linear relationships between variables, GBR leverages multiple weak learners to improve predictive accuracy iteratively. This iterative process helps capture non-linear patterns in the data, leading to more precise grade predictions.The high R-squared value (0.995) indicates that the GBR model explains nearly all the variability in the academic grades, showcasing its robustness and reliability. The low values of MSE and MAE further underscore the model's precision in minimizing prediction errors, making it a valuable tool for educators and policymakers.

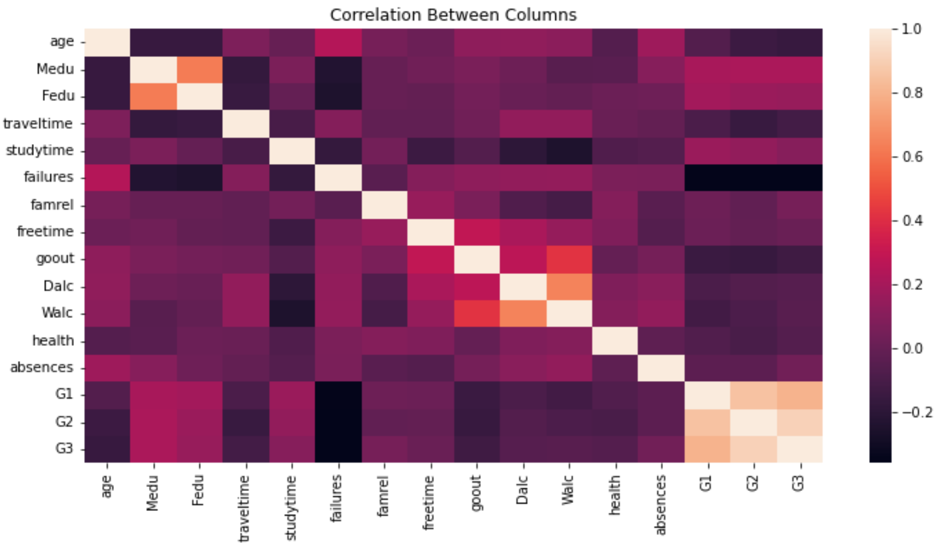

Figure 7. A correlated heatmap matrices for the dataset's parameters

Table 4. The experimental results using the mentioned algorithms.

Algorithm	Training Score	Testing Score	MSE	MAE	MedSE	R2 Score
LR	0.974	0.970	N/A	5.643	1.0305	0.8860
DT	0.997	0.991	1.059	0.314	0.0001	0.990
RF	0.996	0.961	2.080	0.355	0.0922	0.981
GBR	0.999	0.96	0.5068	0.258	.00824	0.995

5. Conclusion and Discussion

This study aimed to enhance the accuracy of academic grade forecasting by employing advanced predictive modeling techniques. We evaluated the performance of Linear Regression (LR), Decision Trees (DT), Random Forest (RF), and Gradient Boosting Regression (GBR) to determine the most effective algorithm for predicting student performance. Our findings demonstrate that Gradient Boosting Regression (GBR) outperforms the other models across all metrics, achieving a training score of 0.999, a testing score of 0.96, an MSE of 0.5068, an MAE of 0.258, and an R^2 of 0.995. While the results of this study are promising, there are some limitations to consider. The dataset used for model training and testing may not capture all the nuances of student performance, such as socio-economic factors, psychological aspects, and extracurricular activities. Future research could incorporate a more diverse set of features to improve the model's generalizability. Additionally, while GBR showed the highest accuracy in this study, it is essential to validate these findings across different educational contexts and datasets. Comparative studies involving other advanced algorithms, such as neural networks or support vector machines, could provide further insights into the most effective approaches for academic grade forecasting.

References

[1] H. Bai, X. Liu, F. Bai, Y. Chen, and R. Pandohie, “Machine Learning Method for High-Dimensional Education Data,” J. Methods Meas. Soc. Sci., vol. 13, no. 1, 2022, doi: 10.2458/jmmss.5396.

[2] S. T. Ahmed, R. S. Al-Hamdani, and M. S. Croock, “EDM preprocessing and hybrid feature selection for improving classification accuracy,” J. Theor. Appl. Inf. Technol., vol. 97, no. 1, pp. 279–289, 2019.

[3] A. Goodarzi, "From Data to Decisions : Mining Data to Forecast Academic Performance," no. April, 2024, doi: 10.13140/RG.2.2.32217.53607.

[4] S. T. Ahmed, R. Al-Hamdani, and M. S. Croock, "Developed third iterative dichotomizer based on feature decisive values for educational data mining," Indones. J. Electr. Eng. Comput. Sci., vol. 18, no. 1, pp. 209–217, 2019, doi: 10.11591/ijeecs.v18.i1.pp209-217.

[5] Cem Özkurt, "Analysis and Evaluation of Factors Influencing Student Success with Explainable Artificial Intelligence Models," pp. 72–78, 2024, doi: 10.56741/esl.v3i03.556.

[6] Hassan, E., Talaat, F. M., Adel, S., Abdelrazek, S., Aziz, A., Nam, Y., & El-Rashidy, N. (2023). Robust Deep Learning Model for Black Fungus Detection Based on Gabor Filter and Transfer Learning. Computer Systems Science & Engineering, 47(2).

[7] Kotsiantis, S. B. (2012). Use of machine learning techniques for educational proposes: a decision support system for forecasting students' grades. Artificial Intelligence Review, 37, 331-344.

[9] Khan, I., Ahmad, A. R., Jabeur, N., & Mahdi, M. N. (2021). An artificial intelligence approach to monitor student performance and devise preventive measures. Smart Learning Environments, 8, 1-18.

[10] Hassan, E., & El-Rashidy, N. M Talaa, F.(2022). mask R-CNN models. Nile Journal of Communication and Computer Science, 3(1), 17-27.

[11] Hassan, E., Bhatnagar, R., & Shams, M. Y. (2023, June). Advancing Scientific Research in Computer Science by ChatGPT and LLaMA—A Review. In International Conference on Intelligent Manufacturing and Energy Sustainability (pp. 23-37). Singapore: Springer Nature Singapore.

[12] de Souza Zanirato Maia, J., Bueno, A. P. A., & Sato, J. R. (2023). Applications of Artificial Intelligence Models in Educational Analytics and Decision Making: A Systematic Review. *World*, *4*(2), 288-313.

[13] Hassan, E., Shams, M. Y., Hikal, N. A., & Elmougy, S. (2023). The effect of choosing optimizer algorithms to improve computer vision tasks: a comparative study. Multimedia Tools and Applications, 82(11), 16591-16633.

[14] Badal, Y. T., & Sungkur, R. K. (2023). Predictive modelling and analytics of students' grades using machine learning algorithms. Education and information technologies, 28(3), 3027-3057.

[15] Olaleye, T. O., & Vincent, O. R. (2020, March). A predictive model for students' performance and risk level indicators using machine learning. In 2020 International Conference in Mathematics, Computer Engineering and Computer Science (ICMCECS) (pp. 1-7). IEEE.

[16] Hussain, S., & Khan, M. Q. (2023). Student-performulator: Predicting students' academic performance at secondary and intermediate level using machine learning. *Annals of data science*, *10*(3), 637-655.

[17] Elbedwehy, S., Hassan, E., Saber, A., & Elmonier, R. (2024). Integrating neural networks with advanced optimization techniques for accurate kidney disease diagnosis. *Scientific Reports*, *14*(1), 21740.

[18] Shams, M. Y., Abd El-Hafeez, T., & Hassan, E. (2024). Acoustic data detection in large-scale emergency vehicle sirens and road noise dataset. *Expert Systems with Applications*, *249*, 123608.

[19] Talaat, F. M., El-Sappagh, S., Alnowaiser, K., & Hassan, E. (2024). Improved prostate cancer diagnosis using a modified ResNet50-based deep learning architecture. *BMC Medical Informatics and Decision Making*, *24*(1), 23.

[20] Hassan, E., Saber, A., & Elbedwehy, S. (2024). Knowledge distillation model for Acute Lymphoblastic Leukemia Detection: Exploring the impact of nesterov-accelerated adaptive moment estimation optimizer. *Biomedical Signal Processing and Control*, *94*, 106246.

[21] Alnowaiser, K., Saber, A., Hassan, E., & Awad, W. A. (2024). An optimized model based on adaptive convolutional neural network and grey wolf algorithm for breast cancer diagnosis. *PloS one*, *19*(8), e0304868.

[22] Shoaib, M., Sayed, N., Singh, J., Shafi, J., Khan, S., & Ali, F. (2024). AI student success predictor: Enhancing personalized learning in campus management systems. *Computers in Human Behavior*, *158*, 108301.

[23] Baniata, L. H., Kang, S., Alsharaiah, M. A., & Baniata, M. H. (2024). Advanced Deep Learning Model for Predicting the Academic Performances of Students in Educational Institutions. *Applied Sciences*, *14*(5), 1963.