Journal of Intelligent Decision Support System (IDSS) 8 (3) (2025), pp. 188-197

Published by: Institute of Computer Science (IOCS)

Journal of Intelligent Decision Support System (IDSS)

Journal homepage: www.idss.iocspublisher.org

Comparison of algorithm performance, Random Forest Regression, SVR, and Gradient Boosting in predicting academic grades based on student lifestyle

Danny sihombing

Informatika, Fakultas Teknik, Universitas HKBP Nommensen Medan, Medan, Indonesia

Article Info

Article history:

Received Sep 03, 2025
Revised Sep 25, 2025
Accepted Sep 29, 2025

Keywords:

Academic Performance
Gradient Boosting
Random Forest
Student Lifestyle
SVR

ABSTRACT

This study examines the effectiveness of three machine learning algorithms—Random Forest Regression, Support Vector Regression, and Gradient Boosting—in predicting students' academic grades based on lifestyle-related factors including study hours, sleep duration, social interaction, physical activity, and stress levels. Employing a quantitative experimental approach, model performance was evaluated using R^2, MSE, RMSE, and MAE, while SHAP analysis was applied to interpret feature importance. The results show that all models achieved reasonable predictive accuracy, with Gradient Boosting consistently outperforming the others across all metrics. Study duration was identified as the most influential predictor, whereas stress level and gender had minimal impact. These findings emphasize the importance of non-academic lifestyle factors in predicting academic achievement and provide insights for the development of data-driven, personalized decision support systems in education.

Corresponding Author:

Danny sihombing,
Informatika, Fakultas Teknik,
Universitas HKBP Nommensen Medan,
Jl. Sutomo No.4A, Perintis, Kec. Medan Tim., Kota Medan, Sumatera Utara 20235, Indonesia.
Email: danny@uhn.ac.id

Introduction

In the modern educational context, academic achievement remains the main benchmark in assessing both student success and the effectiveness of educational institutions. While cognitive ability and teaching quality have long been recognized as key determinants of learning success, non-academic factors particularly students' daily lifestyle habits are increasingly gaining attention as elements that contribute to academic achievement (Han et al., 2022; Huddar, 2021). Variables such as study duration, sleep quality, social interaction, physical activity, and stress levels form a complex behavioral ecosystem in a student's academic journey (Alj & Bouayad, 2024; Lisnyj et al., 2021). Unfortunately, these factors are often poorly accommodated in existing educational prediction models. With the increasing availability of student behavioral data and advances in computing technology, machine learning algorithms have emerged as a powerful approach in modeling and forecasting academic outcomes based on multiple dimensions of data (Onyema et al., 2022). To optimally harness this potential, a rigorous and data driven approach is needed to analyze the extent to which lifestyle affects learning achievement. This is important so that education stakeholders can design personalized and preventive interventions. This

research comes within that framework, with the aim of bridging student behavior data and academic predictions through a systematic and scalable machine learning approach.

Despite various efforts to improve the quality of learning, the prediction of student academic achievement still relies heavily on conventional indicators such as previous grades, attendance, or socio-economic background. This approach has not been able to fully capture the dynamics of students' daily behavior, even though daily habits such as study duration, sleep time, physical activity intensity, and stress levels have significant potential to affect academic outcomes. On the other hand, the development of machine learning algorithms provides a great opportunity to integrate multidimensional data to produce more accurate and adaptive predictive models (Wilson & Anwar, 2024). However, there is no clear consensus on which machine learning algorithms are most effective in modeling the relationship between students' lifestyles and their academic performance (Rajendran et al., 2022). The absence of a systematic comparative mapping of the performance of various algorithms, such as Random Forest, Support Vector Machine Regression, and Gradient Boosting leaves important questions regarding the reliability, efficiency, and interpretability of the models in the context of educational prediction. Therefore, it is necessary to conduct an in depth investigation to identify the most optimal algorithm in predicting academic grades based on student lifestyle data in an empirical and measurable manner.

Various previous studies have explored the use of machine learning algorithms in student academic achievement prediction. For example, a study by Doz et al. (2023) showed that Random Forest Regression and Fuzzy Logic has a good regression ability in predicting academic performance based on demographic student data. A study by (Gaftandzhieva et al., 2022) used Gradient Boosting showed that students' online learning activities, including interaction on Moodle and attendance at Zoom lectures, were significantly correlated with their final grades. Meanwhile, a study by (Muhammad et al., 2023) used Support Vector Regressor showed that socioeconomic factors have a significant influence on student learning achievement, with promising results in terms of accuracy. However, most of these studies have not comprehensively considered student lifestyle variables such as sleep habits, stress, and physical activity as predictive determinants. In addition, the comparative approach between algorithms is often conducted without considering the uniformity of experimental design and performance evaluation in the context of multidimensional data of student behavior. Some previous studies have also recommended that future research include more contextual and lifestyle variables to improve the validity of the model. Therefore, this research is directed to address these limitations by proposing a more holistic and balanced experimental framework in comparing the performance of Random Forest, SVR and Gradient Boosting testing the relevance of lifestyle variables in academic prediction quantitatively and empirically.

This research aims to evaluate and compare the performance of three machine learning algorithms namely Random Forest Regression, Support Vector Regressor, and Gradient Boosting in predicting students' academic grades based on daily lifestyle variables, such as study duration, sleep time, social activity, physical activity, and stress level. Through a systematic computational approach, this research is expected to identify the most optimal algorithm in terms of accuracy, efficiency, and generalization ability on student behavioral data. In addition, this research also aims to uncover the relative contribution of each lifestyle variable to the prediction results, so as to provide a deeper understanding of the non-academic factors that influence learning achievement. The findings of this research are expected to serve as a foundation for the development of data-driven decision support systems in educational contexts, as well as contribute to more personalized, predictive, and evidence-based learning strategies.

This research offers an original contribution by integrating student lifestyle dimensions such as study duration, sleep time, social activity, physical activity, and stress level into an academic grade prediction model using a machine learning approach. The novelty lies in testing and directly comparing three popular algorithms, namely Random Forest Regression, Support Vector Regression, and Gradient

Boosting in a unified analysis framework that focuses on non-academic variables. Most previous studies have focused on conventional academic indicators, thus missing the utilization of students' daily behavioral data that is more dynamic and representative of their actual conditions. By presenting a comparative analysis of model performance, this study not only enriches the literature on machine learning-based academic achievement prediction, but also provides practical justification for educational institutions in choosing the most efficient and accurate algorithm to detect potential academic risks early. The results of this study are expected to be used as a scientific basis for more holistic, predictive, and data driven educational decision making.

Although various international studies have examined machine learning-based academic achievement predictions, the integration of lifestyle variables into prediction models is still far from optimal. Most studies still focus on conventional indicators such as test scores, attendance, or demographic factors, thereby ignoring the daily dynamics of students, which are more representative in describing their actual conditions. This gap is even more significant in the Indonesian context, where vocational and general education systems often face challenges related to limited teacher resources for personal monitoring and increasing academic pressure on students. Therefore, this study not only offers an academic contribution through a comparison of Random Forest, Support Vector Regression, and Gradient Boosting algorithms in a uniform experimental framework, but also provides practical urgency. By utilizing student lifestyle data as the basis for prediction, the results of this study have the potential to assist schools, teachers, and policymakers in Indonesia in designing earlier, more personalized, and evidence-based interventions to improve learning achievement while supporting student well-being.

Method

Research Design

This research uses a quantitative approach with a comparative experimental design. The aim is to compare the performance of three machine learning algorithms namely Random Forest Regression, Support Vector Regression, and Gradient Boosting in predicting student academic grades based on lifestyle data. This research was conducted in several systematic stages, including data collection, preprocessing, modeling, model performance evaluation, and interpretation of results. The entire experimental process was conducted in a controlled computing environment using Python Google Collaboratory software with scikit-learn, matplotlib, sciborn, numpy and pandas libraries.

Data Collection

The dataset used in this study was sourced from Kaggle (*Student Lifestyle Dataset*) and collected via Google Forms from college students. It contains variables on study habits, sleep, physical activity, and social interaction, as summarized in Table 1.

Table 1. Pieces of Research Dataset

Student_ ID	Study_ Hours_ Per_Day	Extracurr icular_ Hours_ Per_Day	Sleep_ Hours_Pe r_Day	Social_ Hours_ Per_Day	Physical_ Activity_ Hours_ Per_Day	Stress_ Level	Gender	Grades
1	6.9	3.8	8.7	2.8	1.8	Moderate	Male	7.48
2	5.3	3.5	8.0	4.2	3.0	Low	Female	6.88
3	5.1	3.9	9.2	1.2	4.6	Low	Male	6.68
4	6.5	2.1	7.2	1.7	6.5	Moderate	Male	7.2
1	6.9	3.8	8.7	2.8	1.8	Moderate	Male	7.48

Data Pre-Processing

In the pre-processing stage, minimal changes were needed since the dataset was already clean and well-structured. However, the categorical columns Stress_Level and Gender were label-encoded to convert them into numerical format for compatibility with machine learning algorithms.

Split Data

Before model training, the dataset was split into 80% training data (1,600 records) and 20% test data (400 records). The training data was then used to train three regression algorithms—Random Forest Regressor, Support Vector Regression (SVR), and Gradient Boosting Regressor—to learn predictive patterns that would later be applied to the test data for evaluation.

Random Forest Regression

Random Forest Regression is an ensemble learning-based machine learning algorithm used for regression tasks, that is predicting a continuous target variable (Bakır et al., 2024; Natras et al., 2022). Scientifically, it is an extension of bagging (bootstrap aggregating) that combines results from a number of decision trees to improve prediction accuracy and reduce model variance (Becker et al., 2023; Syam & Kaul, 2021).

From the original dataset $\boldsymbol{D} = \{\{\boldsymbol{x_1}, \boldsymbol{y_1}, \dots, (\boldsymbol{x_N}, \boldsymbol{y_N})\}$, a bootstrap dataset B is formed (by random sampling with returns).

For each bootstrap dataset:

a. Train a decision tree $\boldsymbol{T_b}$.
b. At each node, select a random subset of features, then find the best split.

Each tree produces a prediction $\boldsymbol{T_b}(\boldsymbol{x})$.
The final prediction for the input $(\boldsymbol{x})$ is calculated as the average of all tree predictions:

$$\boldsymbol{y}(\boldsymbol{x}) = \frac{\mathbf{1}}{\boldsymbol{B}} \sum_{\boldsymbol{b}=\mathbf{1}}^{\boldsymbol{B}} \boldsymbol{T_b}(\boldsymbol{x}) \quad (1)$$

Super Vector Regression

Support Vector Regression (SVR) is a regression method that predicts continuous values by finding the best function that stays within a certain margin of error tolerance (epsilon) while keeping the model simple and not overfitting (Montesinos López et al., 2022; Zhang & O'Donnell, 2020). With the help of kernel functions, SVR is also capable of handling non-linear relationships, although its performance may degrade on very large datasets and is sensitive to parameter selection (Najafzadeh & Niazmardi, 2021; Otchere et al., 2021).

SVR aims to find the regression function from the shape:

$$\boldsymbol{f}(\boldsymbol{x}) = \boldsymbol{w}^{\boldsymbol{T}} \emptyset(\boldsymbol{x}) + \boldsymbol{b} \quad (2)$$

$\boldsymbol{where}\ \emptyset(\boldsymbol{x})$: (Non-linear) transformation of input features to a higher dimensional feature space, $\boldsymbol{w}$: Weight vector, $\boldsymbol{b}$: Bias/intercept. The goal is to make $\boldsymbol{f}(\boldsymbol{x})$ as close as possible to $\boldsymbol{y}$, but with an error tolerance of $\boldsymbol{\varepsilon}$ (epsilon-insensitive zone).

Gradient Boosting

Gradient Boosting Regressor is a machine learning algorithm for continuous value prediction (regression) that builds models incrementally by sequentially merging many weak decision trees, where each new tree is trained to correct the prediction error of the previous model by minimizing a loss function using a gradient approximation (Khamis et al., 2024).

Gradient Boosting models the prediction function as a summation of weak models (usually shallow decision trees):

$$\boldsymbol{F(x) = \sum_{m=1}^{M} y_m h_m(x)} \quad (3)$$

Where: $F(x)$: final prediction function, $h_m(x)$: m-th weak model (usually tree), γ_m: step size or learning rate, M: number of iterations/tree

Regression

After training the data with the three regression algorithms, the next step is to evaluate the performance of each model on the test data by calculating regression metrics, namely the coefficient of determination (R^2), mean squared error (MSE), root mean squared error (RMSE), and mean absolute error (MAE).

Evaluation Metrics

a. Mean Squared Error (MSE)

Measures the average of the squared difference between the actual and predicted values (Hodson et al., 2021):

$$\boldsymbol{MSE = \frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2} \quad (5)$$

Sensitive to outliers as errors are squared.

b. Root Mean Squared Error (RMSE)

Is the root of the MSE, so it returns to the original units of the target (Hodson, 2022):

$$\text{RMSE}=\sqrt{\boldsymbol{MSE}} \quad (6)$$

Eases interpretation as it is on the same scale as $\boldsymbol{y}$.

c. Mean Absolute Error (MAE)

Measures the average of the absolute value of the difference between the actual and predicted values (Robeson & Willmott, 2023):

$$\boldsymbol{MAE = \frac{1}{N}\sum_{i=1}^{N}|y_i - \hat{y}_i|} \quad (7)$$

More resistant to outliers than MSE.

d. Coefficient Determination ($\boldsymbol{R^2}$)

Assesses how much variation in the target data can be explained by the model (Avdeef, 2021):

$$\boldsymbol{R^2 = 1 - \frac{\sum_{i=1}^{N}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{N}(y_i - \bar{y}_i)^2}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2} \quad (8)$$

$$\boldsymbol{\bar{y} = \frac{1}{N}\sum_{i=1}^{N} y_i} \quad (9)$$

where: $\boldsymbol{R^2 = 1}$, means the prediction is perfect, $\boldsymbol{R^2 = 0}$, means that the model is no better than the average, It can be negative if the model is very bad.

e. SHAP

SHAP (SHapley Additive exPlanations) is a machine learning model interpretability method used to explain the contribution of each feature to the prediction results of a model (Ekanayake et al., 2022). SHAP is based on the Shapley value theory from cooperative game theory, which fairly distributes the "contribution" of each feature by considering all possible combinations of features (Liu et al., 2024; Veeramsetty, 2021).

The SHAP value for a feature *i* is defined as:

$$\emptyset_i = \mathbf{1} \sum_{S \subseteq N \setminus \{i\}} \frac{|S|!.\,(|N| - |S| - \mathbf{1})!}{|N|!} [f(S \cup \{i\}) - f(S)] \quad (10)$$

Where: $\emptyset_i$ is the SHAP value for feature iii, representing its contribution to the model prediction, N is the set of all features, S is a subset of features not containing I, $f(S)$ is the model prediction using only the features in subset S, The expression $f(S \cup \{i\}) - f(S)$ measures the marginal contribution of feature i when added to subset S, The term $\frac{|S|!.(|N|-|S|-1)!}{|N|!}$ is a weighting factor that ensures fairness by averaging over all possible orderings of features.

Results and Discussions

The data training process carried out on the train data of 1600 records using the three regression algorithms and the regression applied to the test data of 400 records resulted in accuracy as presented in Figure 1 below.

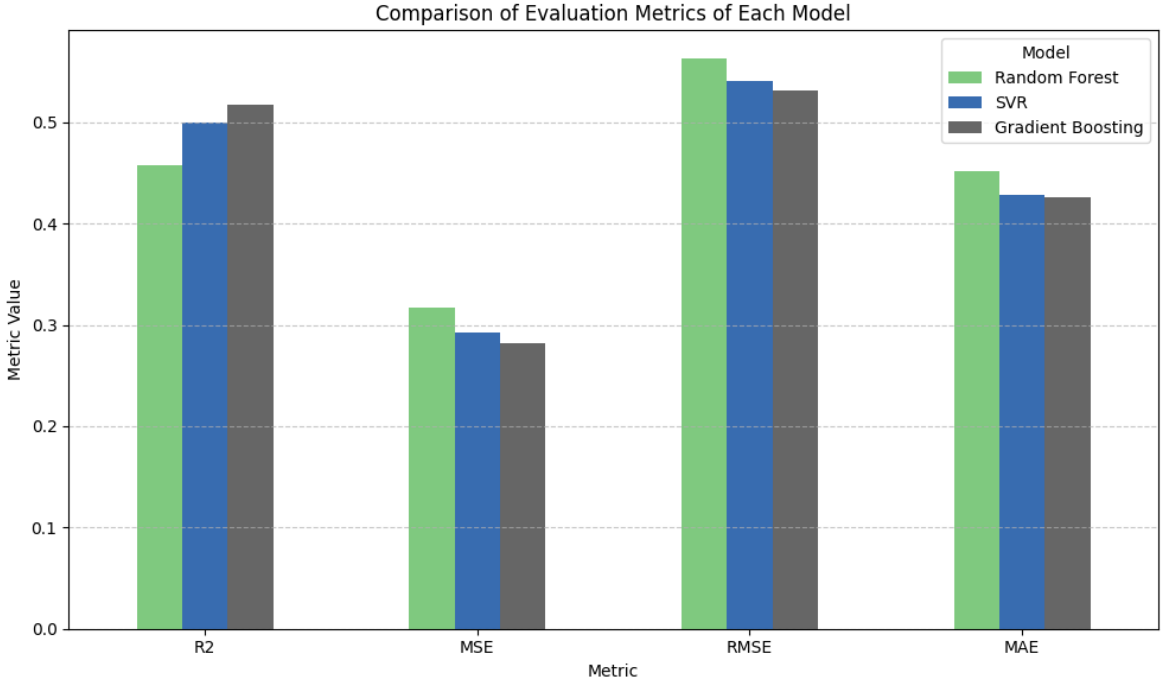

Figure 1. Metric Value Comparison of Each Model

Figure 1 shows that the Gradient Boosting Regressor achieves the highest R^2 score, indicating the best ability to explain the variance in the target variable compared to Random Forest and SVR. While all models show comparable performance, Gradient Boosting also attains the lowest MSE and MAE, suggesting that its predictions are generally closer to the actual values. Despite SVR having a slightly lower RMSE than Gradient Boosting, the difference is minimal. Therefore, in terms of both predictive

accuracy and error reduction, Gradient Boosting performs most consistently across all evaluation metrics.

Furthermore, to find out which features have the most impact on the regression of grades value, we can look at two graphs, namely SHAP.

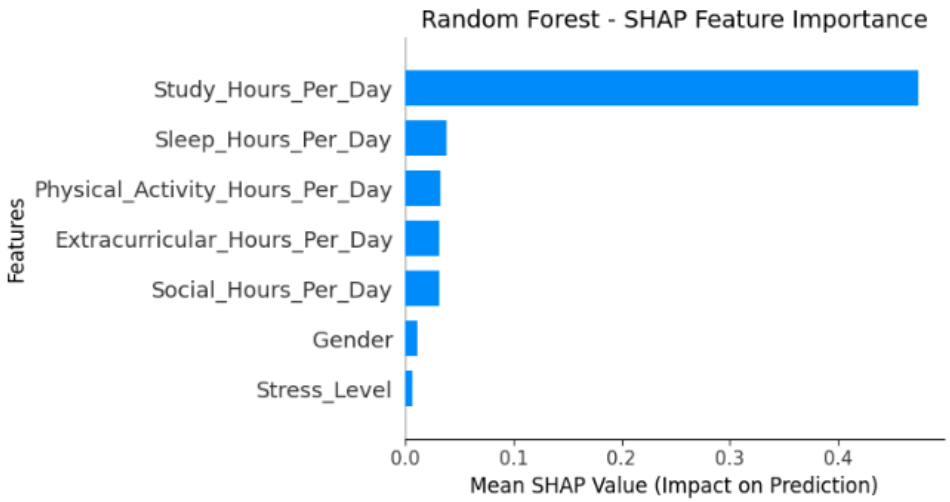

Figure 2. Feature Importance Random Forest Model

Figure 2 illustrates the SHAP feature importance for the Random Forest Regression that the feature Study_Hours_Per_Day has the largest impact on the model's prediction, as indicated by the highest mean SHAP value. This suggests that the amount of time a student spends studying per day is the most influential factor in predicting the target variable (e.g., performance or academic outcome) within the Random Forest model. Other features have relatively low but noticeable contributions.

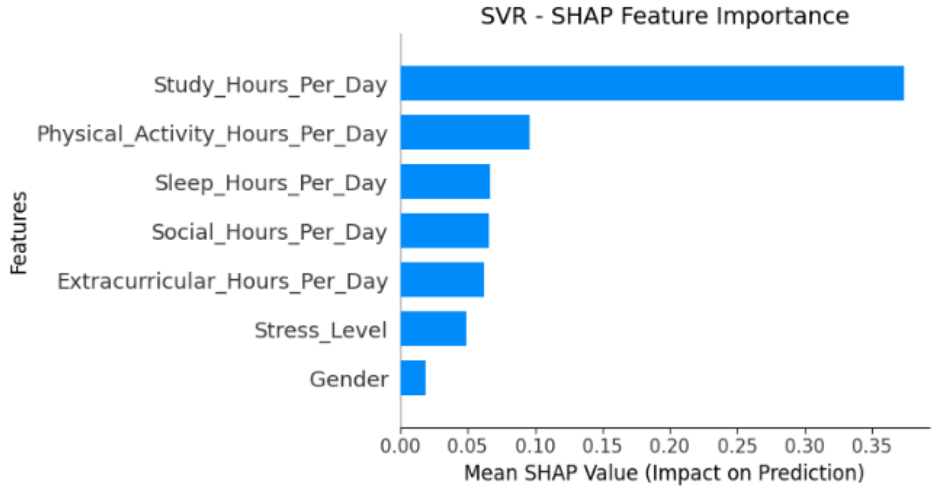

Figure 3. Feature Importance SVR Model

Figure 3 illustrates the SHAP feature importance for the Support Vector Regression (SVR) model, highlighting that Study_Hours_Per_Day is the most influential predictor in determining the model's output. Unlike the Random Forest model, where one feature dominates, the SVR model distributes importance more evenly across several features. Variables such as Physical_Activity_Hours_Per_Day, Sleep_Hours_Per_Day, Social_Hours_Per_Day, and Extracurricular_Hours_Per_Day show moderate contributions, indicating that SVR captures more subtle

and multifactorial relationships in the data. In contrast, Gender and Stress_Level have minimal impact, suggesting limited relevance to the target variable. This balanced contribution reflects SVR's ability to model complex interactions, provided proper feature scaling and parameter tuning are applied.

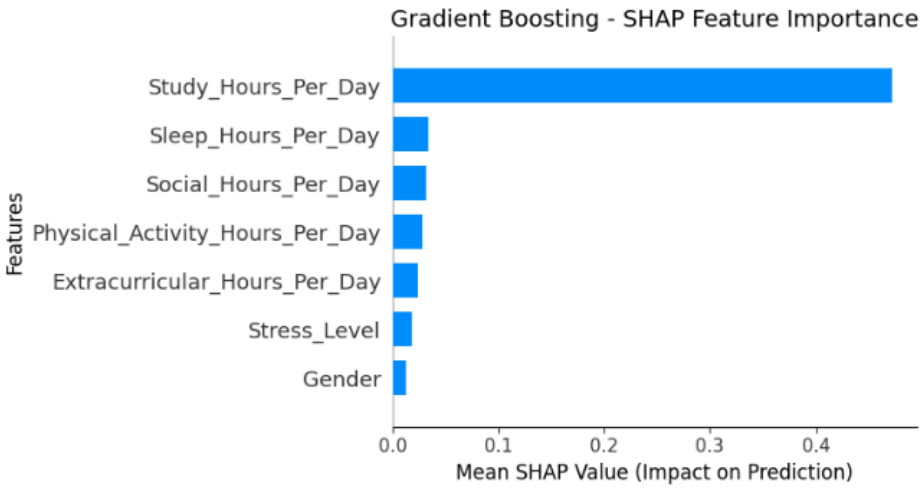

Figure 4. Feature Importance Gradient Boosting Model

Figure 4 displays the SHAP feature importance for the Gradient Boosting Regressor model, where Study_Hours_Per_Day clearly stands out as the most significant feature influencing the model's predictions. Its SHAP value far exceeds that of all other features, indicating that study time has the strongest impact on the target variable. Meanwhile, features such as Sleep_Hours_Per_Day, Social_Hours_Per_Day, Physical_Activity_Hours_Per_Day, and Extracurricular_Hours_Per_Day contribute with relatively minor yet comparable importance. Stress_Level and Gender again show minimal influence, consistent with the patterns observed in the other models. This suggests that while Gradient Boosting captures some variation across lifestyle factors, it strongly prioritizes academic engagement as the key predictive factor.

Conclusions

Based on the experimental results, all three regression algorithms—Random Forest Regressor, Support Vector Regression (SVR), and Gradient Boosting Regressor—were able to model student grade predictions with reasonable accuracy. Among them, Gradient Boosting consistently outperformed the others across all evaluation metrics (R^2, MSE, RMSE, and MAE). SHAP analysis further revealed that Study_Hours_Per_Day was the most influential feature in all models, confirming the strong correlation between study duration and academic performance. Meanwhile, features such as sleep, social interaction, and physical activity contributed moderately, while gender and stress level showed minimal impact. Practically, these findings highlight the potential of machine learning-based prediction models as decision support tools for educational institutions. Schools and teachers can utilize such models to identify students at academic risk early and design data-driven interventions, such as personalized study plans, time management training, or lifestyle adjustment programs, to improve learning outcomes. However, this research is limited by the scope of variables considered, the relatively small sample size, and the absence of longitudinal data that could capture changes in student behavior over time. Future studies should expand the dataset, include additional psychological or environmental factors, and explore hybrid or deep learning models to enhance predictive accuracy and generalizability. By addressing these limitations, subsequent research can provide more robust insights, thereby strengthening the contribution of machine learning approaches to evidence-based educational policy and practice.

References

Alj, Z., & Bouayad, A. (2024). Multidimensional determinants of academic performance: Insights from undergraduate students in Moroccan universities. *JOTSE: Journal of Technology and Science Education*, *14*(2), 607–621.

Avdeef, A. (2021). Do you know your r2? *ADMET and DMPK*, *9*(1), 69–74.

Bakır, R., Orak, C., & Yüksel, A. (2024). Optimizing hydrogen evolution prediction: A unified approach using random forests, lightGBM, and Bagging Regressor ensemble model. *International Journal of Hydrogen Energy*, *67*, 101–110.

Becker, T., Rousseau, A.-J., Geubbelmans, M., Burzykowski, T., & Valkenborg, D. (2023). Decision trees and random forests. *American Journal of Orthodontics and Dentofacial Orthopedics*, *164*(6), 894–897.

Doz, D., Cotič, M., & Felda, D. (2023). Random forest regression in predicting students' achievements and fuzzy grades. *Mathematics*, *11*(19), 4129.

Ekanayake, I. U., Meddage, D. P. P., & Rathnayake, U. (2022). A novel approach to explain the black-box nature of machine learning in compressive strength predictions of concrete using Shapley additive explanations (SHAP). *Case Studies in Construction Materials*, *16*, e01059.

Gaftandzhieva, S., Talukder, A., Gohain, N., Hussain, S., Theodorou, P., Salal, Y. K., & Doneva, R. (2022). Exploring Online Activities to Predict the Final Grade of Student. *Mathematics*, *10*(20), 1–20. https://doi.org/10.3390/math10203758

Han, C., Farruggia, S. P., & Solomon, B. J. (2022). Effects of high school students' noncognitive factors on their success at college. *Studies in Higher Education*, *47*(3), 572–586.

Hodson, T. O. (2022). Root mean square error (RMSE) or mean absolute error (MAE): When to use them or not. *Geoscientific Model Development Discussions*, *2022*, 1–10.

Hodson, T. O., Over, T. M., & Foks, S. S. (2021). Mean squared error, deconstructed. *Journal of Advances in Modeling Earth Systems*, *13*(12), e2021MS002681.

Huddar, N. M. (2021). *A Study of Life Style of Low Achievers with Regard to Their Academic Life, Non Academic Life, Parents Involvement and Self Regulating Learning Habbit*.

Khamis, G. S. M., Mohammed, Z. M. S., Alanazi, S. M., Mahmoud, A. F. A., Abdalla, F. A., & Bkheet, S. A. (2024). Prediction of Myocardial Infarction Complications using Gradient Boosting. *Engineering, Technology & Applied Science Research*, *14*(6), 18550–18556.

Lisnyj, K. T., Pearl, D. L., McWhirter, J. E., & Papadopoulos, A. (2021). Exploration of factors affecting post-secondary students' stress and academic success: Application of the socio-ecological model for health promotion. *International Journal of Environmental Research and Public Health*, *18*(7), 3779.

Liu, Y., Fu, Y., Peng, Y., & Ming, J. (2024). Clinical decision support tool for breast cancer recurrence prediction using SHAP value in cooperative game theory. *Heliyon*, *10*(2).

Montesinos López, O. A., Montesinos López, A., & Crossa, J. (2022). Support vector machines and support vector regression. In *Multivariate Statistical Machine Learning Methods for Genomic Prediction* (pp. 337–378). Springer.

Muhammad, Y., Hassan, M. A., Almotairi, S., Farooq, K., Granelli, F., & Strážovská, Ľ. (2023). The Role of Socioeconomic Factors in Improving the Performance of Students Based on Intelligent Computational Approaches. *Electronics (Switzerland)*, *12*(9), 1–18. https://doi.org/10.3390/electronics12091982

Najafzadeh, M., & Niazmardi, S. (2021). A novel multiple-kernel support vector regression algorithm for estimation of water quality parameters. *Natural Resources Research*, *30*(5), 3761–3775.

Natras, R., Soja, B., & Schmidt, M. (2022). Ensemble machine learning of random forest, AdaBoost and XGBoost for vertical total electron content forecasting. *Remote Sensing*, *14*(15), 3547.

Onyema, E. M., Almuzaini, K. K., Onu, F. U., Verma, D., Gregory, U. S., Puttaramaiah, M., & Afriyie, R. K. (2022). Prospects and challenges of using machine learning for academic forecasting. *Computational Intelligence and Neuroscience*, *2022*(1), 5624475.

Otchere, D. A., Ganat, T. O. A., Gholami, R., & Ridha, S. (2021). Application of supervised machine learning paradigms in the prediction of petroleum reservoir properties: Comparative analysis of ANN and SVM models. *Journal of Petroleum Science and Engineering*, *200*, 108182.

Rajendran, S., Chamundeswari, S., & Sinha, A. A. (2022). Predicting the academic performance of middle-and high-school students using machine learning algorithms. *Social Sciences & Humanities Open*, *6*(1), 100357.

Robeson, S. M., & Willmott, C. J. (2023). Decomposition of the mean absolute error (MAE) into systematic and unsystematic components. *PloS One*, *18*(2), e0279774.

Syam, N., & Kaul, R. (2021). Random forest, bagging, and boosting of decision trees. In *Machine Learning and Artificial Intelligence in Marketing and Sales: Essential Reference for Practitioners and Data Scientists* (pp. 139–182). Emerald Publishing Limited.

Veeramsetty, V. (2021). Shapley value cooperative game theory-based locational marginal price computation for

loss and emission reduction. *Protection and Control of Modern Power Systems*, *6*(4), 1–11.
Wilson, A., & Anwar, M. R. (2024). The Future of Adaptive Machine Learning Algorithms in High-Dimensional Data Processing. *International Transactions on Artificial Intelligence*, *3*(1), 97–107.
Zhang, F., & O'Donnell, L. J. (2020). Support vector regression. In *Machine learning* (pp. 123–140). Elsevier.