Proceeding of the International Conference on Management, Entrepreneurship, and Business
Volume 2 Nomor 1 Tahun 2025
e-ISSN: XXXX-XXX; p-ISSN: XXXX-XXX, Hal 72 - 83
DOI: https://doi.org/xx.xxxx
Available online at: https://xxxx.xxx

Balancing Time and Success: A Machine Learning Model for GPA Prediction

Huy Hoang Doan[1], Weishen Wu[2]

[1]College of Management, Da-Yeh University, 168 University Rd, Dacun Township, Changhua County, 51591, Taiwan

[2]Department of Accounting and Information Management, Da-Yeh University, 168 University Rd, Dacun Township, Changhua County, 51591, Taiwan

Abstract:

This study explores the application of machine learning to predict students' GPA based on behavioral and time-related factors, including study hours, extracurricular activities, sleep, social interactions, and physical activity. Seven regression algorithms were employed to evaluate predictive accuracy using metrics such as MAE, RMSE, and R^2 Among these, Regularized Linear Regression demonstrated the highest accuracy and interpretability, highlighting its suitability for this dataset. The findings emphasize the potential of machine learning in identifying key predictors of academic performance and offer practical applications for personalized academic advising and time management. This research provides a data-driven framework to support students and educators in optimizing learning outcomes.

Keywords: Machine Learning, Predictive Analytics, Education, GPA

1. Introduction

In an era where academic success is generally measured by a single metric Grade Point Average (GPA), the question arises: how well do we understand the many elements that affect this essential measure of student performance? In light of this, as educational institutions adopt data-driven methodologies, machine learning may help explain GPA prediction. Specifically, study hours, sleep quality, and physical exercise can affect academic achievement; therefore, it's important to understand how they interact. GPA also affects future academic and employment chances, so it is crucial to use modern analytical methods to better evaluate student success in a continuously changing educational landscape.

Thus, this study emphasises the need for using machine learning to predict academic success, specifically GPA. Understanding GPA variables is significant as schools use quantitative metrics to evaluate student performance. Machine learning lets educators and academics find patterns and correlations in complicated datasets that statistical methods may miss. For instance, Iqbal (2017) emphasises the expanding use of machine learning to predict student grades and its potential to improve educational outcomes through data-driven insights. In addition, Thorat (2023) uses machine learning to identify early student success indicators, shows that incorporating behavioural factors like study habits and sleep patterns into predictive models can improve academic performance understanding. As a primary factor of academic performance and future possibilities, GPA is important to academic and career success. Indeed, high GPAs are linked to

Received: August 29, 2023; Accepted: November 22, 2023; Published: Desember 28, 2024
*Corresponding author, e-mail address

greater career prospects, higher earnings, and admission to top graduate programmes (Maia et al., 2023). GPA is also used by schools to determine scholarships and academic honours, which in turn, motivates and engages students. Therefore, GPA issues must be understood by educators and politicians in order to improve student achievements and provide equitable access to opportunities. In effect, accurate GPA prediction allows institutions to target interventions for at-risk students, thus creating a more supportive educational environment (Chen & Cui, 2020). Consequently, this study uses advanced analytical methods to propose solutions to improve academic performance and help students reach their goals.

Understanding behavioural characteristics like study hours, sleep quality and physical activity is essential for predicting academic achievement, especially GPA. Indeed, research has demonstrated that these characteristics are interconnected and strongly affect students' cognitive functioning and academic success. For example, Ali et al. (2013) note that sleep patterns might reduce academic performance and that behavioural interventions like exercise and stress management can help students with sleep issues. Similarly, Al-Khani et al. (2019) also identified a two-way association between sleep quality and physical activity, suggesting that poor sleep typically leads to physical inactivity, which can damage academic performance. Additionally, Putra et al. (2023) also note that poor sleep increases stress and hinders learning. This emphasising the necessity to examine at these psychological factors when creating GPA prediction models. machine learning in predictive analytics for education can significantly change our understanding of academic performance. By leveraging modern algorithms and data analysis, machine learning models can reveal intricate correlations between factors affecting GPA, such as study hours, extracurricular activities, and physical health (Zúñiga-Prado et al., 2023). According to Maia et al. (2023) machine learning algorithms make more accurate predictions and insights into student behaviour than statistical methods. This capability, therefore, helps identify problematic students and develop personalised learning solutions (Singh et al., 2023). By doing so, teachers and researchers can better understand academic performance and improve student success by understanding the complicated relationship between study habits, sleep, and physical exercise.

From these aspects, the main research question addressed in this study is: "Which machine learning models accurately predict GPA using time allocation factors like study hours, extracurricular activities, and physical activity?". Educational analytics can use academic performance prediction to customise their support and improve student outcomes.

2. Contributions

The framework for forecasting GPA using seven machine learning algorithms is the study's main contribution. By employing advanced computational techniques, this methodology aims to improve GPA forecasts, which in turn, is essential for identifying at-risk students and intervening quickly. Notably, Yağcı (2022) showed that data mining and learning analytics can predict student performance in educational contexts. Building on these findings, by integrating many algorithms, this study provides a rigorous framework for analysing multifactorial influences on GPA, including behavioural characteristics such as study hours and sleep patterns.

Besides framework construction, a comparative examination of model performance will be conducted using metrics like MSE, RMSE, MAE, and R^2. These parameters are necessary to assess this study's machine learning models' predictive accuracy. As previous research indicates, these performance metrics have been shown to be crucial for evaluating prediction models in education (Nascimento et al., 2018). By systematically comparing algorithms, this study aims to find the best GPA predictors, thereby advancing educational analytics.

Furthermore, this investigation also identified GPA affecting characteristics, specifically study hours and sleep patterns. Understanding these elements is essential for creating focused student performance plans. Recent research has demonstrated that study time and sleep quality improve academic performance (Mak et al., 2012; Urrila et al., 2017; Perotta et al., 2021). For instance, Creswell et al. (2023; Hershner (2020) discovered that nightly sleep duration significantly predicts college students' GPA. This underscoring the relevance of sleep as a behavioural element in academic achievement (Ursache et al., 2021). Thus, this study seeks to identify these crucial elements to inform educational methods and help students succeed academically.

The final purpose of this research is to give instructors and students actionable knowledge. By translating predictive model findings into practical recommendations, this study empowers instructors and students to make informed academic performance-enhancing decisions. Moreover, the analysis can help educators construct student-specific interventions and teach them good study habits and time management. As highlighted by Yang & Xiang (2024) peer and instructor academic support is vital to student engagement and achievement. Therefore, this study aims to create a supportive educational atmosphere that fosters academic performance as well as student well-being, going beyond just GPA prediction.

3. Background of the Study

Grade Point Average (GPA) is a key indication of academic success and future prospects. educational institutions utilise GPA to evaluate student progress, determine scholarship eligibility and assess preparation for advanced studies or employment. research indicates that a high GPA is linked to greater career opportunities and higher earnings. This highlights its importance in shaping students' academic and professional paths (Negru-Subtirica et al., 2019). Furthermore, GPA might impact students' self-esteem and motivation, as it often reflects their commitment to education and societal norms. However, as GPA is a single measure of accomplishment, it is difficult to appreciate the multiple factors that affect academic performance. Significantly, stress, burnout, and extracurricular involvement might affect GPA, which makes complicates the interpretation (Shadid et al., 2020; Guang et al., 2021).

Despite its importance, the difficulties in identifying GPA effects suggest a more nuanced approach to academic assessment. In fact, according to several research studies, social comparison, mental health, and extracurricular activities can affect GPA and academic achievement (Hakami, 2021; Hu & Cheung, 2021). Shadid et al. (2020) observed that students with lower GPAs are more stressed and less likely to participate in extracurriculars. This can worsen academic issues. Therefore, a broader view of student performance is needed because GPA and other factors like socioeconomic position and culture might affect academic achievements (Zeng, 2023; Hu & Cheung, 2021) while GPA remains an important educational indicator, it is important to evaluate its many determinants in order to develop better methods to enhance student achievement and well-being.

Machine learning is preferred over statistical methods in educational analytics. Indeed, machine learning algorithms can analyse massive volumes of data and find intricate patterns that normal analysts may miss. As a result, this provides more precise student performance projections and academic success criteria detection. Tjahyaningtijas (2023) notes that machine learning allows for student-data-driven customisation of educational materials and instructional approaches which improves learning. Additionally, machine learning can adapt and improve as more data is collected, which makes it a useful tool for educators looking to enhance student results. Unlike linear assumptions and missing complex data linkages, machine learning offers a more dynamic way to study educational phenomena (Krishna, 2021). Moreover, recent educational studies have shown that machine learning can revolutionise student data analysis and interpretation. Krishna (2021)stresses that machine learning can predict student achievement using grades and psychological assessments. Pandian (2024) literature review shows that machine learning algorithms targeting distinct educational needs are becoming more common in student performance predicting. This rising shows that machine learning transforms educational analytics by improving predictive accuracy and informing educational institution decision-making. As the discipline evolves, machine learning will provide greater insights into student behaviour and performance, ultimately creating a more helpful and effective learning environment (Nelyub, 2023).

4. Materials and Methods

4.1 Proposed Framework

The proposed framework presents a structured machine learning GPA prediction process. In the first step, prepare the dataset for analysis. Next, machine learning techniques are consistent and compatible after careful data pre-processing and feature encoding. seven machine learning models are built and assessed to predict GPA using MSE and R². Finally, model performance comparisons determine the best and most efficient algorithm (Fig 1). This structured process systematically explores time allocation elements and GPA correlations, ensuring robust model evaluation.

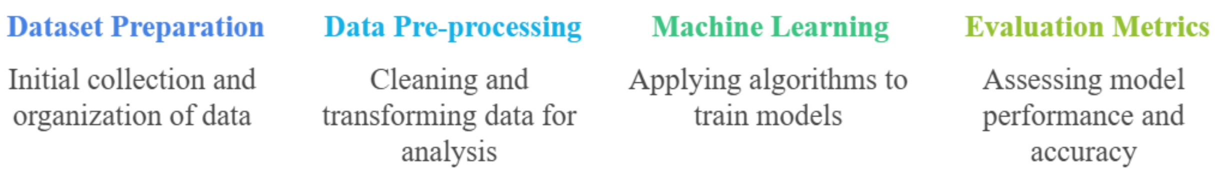

Figure 1. Proposed framework

4.2 Dataset

The dataset provides detailed insights on student lifestyle behaviours and GPA. Students self-reported the data through the Google Forms survey. The form was distributed to students in India and other South Asian nations, regions where the CGPA system is popular. This aspect makes the dataset relevant to various regions' educational environments. The 2,000 Kaggle records document everyday behaviours, namely study hours, extracurricular activities, sleep, social connections, and physical exercise (fig 2). Notably, study and sleep patterns also affect stress levels, presenting a multifaceted view of how lifestyle factors affect GPA. This heterogeneous dataset thus allows robust research and identification of academic achievement determinants. these variables constitute the basis for creating and testing machine learning models that predict GPA by integrating behavioural and time-management aspects.

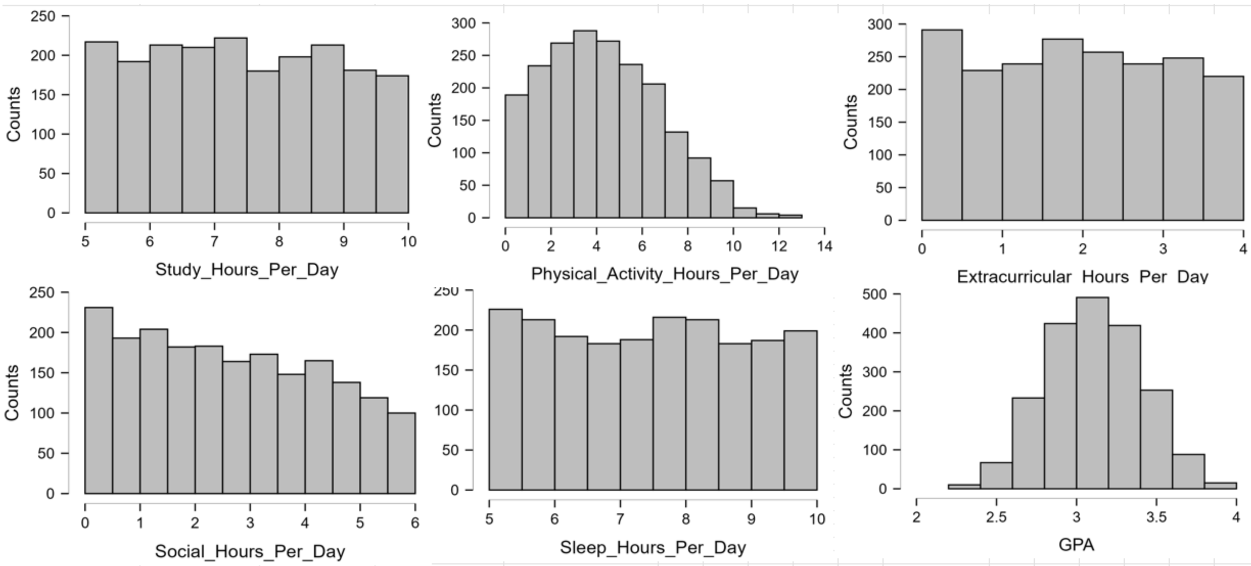

Figure 2. Dataset

4.3 Data Pre-processing

Effective data pre-processing is a important step in any machine learning pipeline as it ensures model dependability. In order to address missing values, we impute missing data with mean values or remove incomplete entries based on their relevance. To avoid distortion caused by anomalies, outliers are discovered and addressed. This approach ensures that data consistency and completeness improve dataset quality, allowing for accurate and relevant analysis. As a result, these procedures create a solid foundation for training machine learning models, enhancing prediction accuracy and performance.

4.4 Features Encoding

Categorical variables like stress level indicators benefit from feature encoding to improve model interpretability and forecast accuracy. Specifically, ordinal encoding incorporates categorical values into rank based scales while one-hot encoding creates binary flags for each category to preserve data granularity. This methodology ensures that machine learning algorithms capture and use stress patterns effectively. After data transformation, methodological adjustments improve the model's academic prediction ability.

4.5 Normalization Method

To achieve consistent feature scaling, normalization adjusts variables like study hours and physical activity to prevent machine learning models from being disproportionately affected. For this purpose, the study uses StandardScaler to standardize variables into Z-scores $= \frac{X - mean(X)}{std(X)}$. This technique ensures that each variable has a mean of zero and a standard deviation of one for consistent model performance and speedier convergence. Thus, the framework optimizes feature compatibility across machine learning methods by standardizing data.

4.6 Machine Learning Algorithms

4.6.1 Boosting Regression

Boosting Regression an ensemble technique that combines the predictions of multiple weak learners to create a strong predictive model. It excels in reducing bias and variance which makes it particularly effective for datasets with complex patterns. Studies have shown that boosting methods can significantly improve prediction accuracy compared to single models. This makes them a valuable choice for educational performance prediction (Shahiri et al., 2015)

4.6.2 Decision Tree Regression

Decision Tree Regression uses a tree-like decision model to explain how different factors affect academic outcomes. It handles category and numerical data, making it useful for educational datasets. Decision trees effectively identify key traits, guiding educators in intervention emphasis (Guarín et al., 2015).

4.6.3 KNN Regression

K-Nearest Neighbors (KNN) Regression predicts outcomes using feature space data points' proximity. KNN works well with little training data because to its simplicity. Setiyorini & Asmono (2020) found that KNN can accurately predict student performance when paired with feature selection.

4.6.4 Neural Network (NN) Regression

Neural Network Regression models complex data interactions using interconnected nodes. This model captures non-linear patterns well, making it suited for educational settings with complex variable interactions. In large datasets, neural networks outperform regression approaches (Gao & Li, 2019).

4.6.5 Random Forest Regression

Random Forest Regression constructs numerous decision trees and averages their predictions. This method improves resilience and lowers overfitting, making it suitable for educational data with intrinsic unpredictability. Educational analytics commonly uses high-dimensional data, which Random Forest handles well (Drzewiecki, 2016).

4.6.6 Regularized Linear Regression (RLR)

Regularized Linear Regression penalizes complex models with Lasso and Ridge regression to prevent overfitting. This approach helps with multicollinearity in predictors, a typical problem in educational datasets. Regularised linear regression provides interpretable results to help educators understand how each variable affects student performance (Shahiri et al., 2015).

4.6.7 SVM Regression

Support Vector Machine (SVM) Regression identifies the best hyperplane to separate data points in high-dimensional space. SVM works when characteristics and outcomes are non-linear. Its robustness and capacity to handle outliers make it a good predictor of academic achievement in varied educational environments (Ananda & Prasetiadi, 2021).

4.8 Performance Metrics

We employed five commonly used evaluation metrics to assess the effectiveness of prediction results and compare different machine-learning techniques. Specifically, we consider the error deviation values of mean squared error (MSE), root mean square error (RMSE), mean absolute error (MAE), and mean absolute percentage error (MAPE). Smaller error values indicate superior model performance. Additionally, we utilize the coefficient of determination (R^2) to measure the explained variance of each model. Greater R-squared values signify a better model for prediction.

5. Experiment

Experiment Setup

The experiment utilized Python as the primary tool for model training and evaluation, supported by standard computational resources. Python provided a user-friendly platform for implementing machine learning algorithms and analyzing results through visualizations and statistical outputs. The experiment was conducted on a mid-range system with sufficient processing power to handle the dataset of 2,000 records efficiently. While not requiring specialized hardware, the setup ensured consistent performance across all models. This setup enabled seamless execution of the data preparation, model training, and evaluation processes, forming a foundation for reliable comparisons.

Splitting Data

The dataset was divided into training (80%) and test (20%) sets to balance model training and performance evaluation. The 80/20 split ensured the model had enough data for learning but also enough to test how well it could be used in other situations. This ratio is widely accepted in machine learning and give a robust framework for assessing model accuracy on unseen data. By maintaining this balance, overfitting was minimized and evaluation results reflected real world performance. This splitting methodology set the stage for fair comparisons among the seven machine learning models tested in the study.

Table 1. Dataset

Variable	Training size of input dataset	Testing size input dataset
Dataset	1600	400

6. Results

Table 2 shows the performance evaluation. It highlights that Regularized Linear Regression and Support Vector Machine Regression are the most effective models for predicting GPA. Both achieved the lowest Mean Squared Error (MSE = 0.04) and Root Mean Squared Error (RMSE = 0.2), along with the lowest Mean Absolute Error (MAE = 0.158) and Mean Absolute Percentage Error (MAPE = 5.19% and 5.18%, respectively). Furthermore, these models demonstrated the highest R2 values with Regularized Linear Regression scoring 0.536 and Support Vector Machine Regression 0.535, indicating superior predictive power and the ability to explain the variance in GPA. While ensemble models like Boosting Regression and Random Forest Regression also performed well, their slightly higher MSE and RMSE suggest they are marginally less precise. In contrast, K-Nearest Neighbors Regression (R2=0.459) and Decision Tree Regression (R2=0.502) struggled to achieve comparable accuracy, likely due to their limitations in handling the dataset's complexity. These findings affirm Regularized Linear Regression and Support Vector Machine Regression as the top choices for robust and interpretable GPA predictions.

Table 2. Machine learning results

	Boosting Regression	Decision Tree Regression	K-Nearest Neighbors Regression	Neural Network Regression	Random Forest Regression	**Regularized Linear Regression**	Support Vector Machine Regression
MSE	0.041	0.043	0.047	0.043	0.042	**0.04**	0.04
RMSE	0.202	0.207	0.217	0.207	0.205	**0.2**	0.2
MAE / MAD	0.159	0.165	0.174	0.163	0.161	**0.158**	0.158
MAPE	5.22%	5.39%	5.69%	5.35%	5.26%	**5.19%**	5.18%
R2	0.532	0.502	0.459	0.502	0.511	**0.536**	0.535

Regularized Linear Regression emerged as the most effective, delivering precise predictions with consistent performance across all metrics. Its simplicity and interpretability further solidify its utility for GPA prediction.

7. Discussion

Regularized Linear Regression outperformed other models due to its simplicity and robustness in handling multicollinearity. This algorithm applies penalties to large coefficients, preventing overfitting while maintaining interpretability. In contrast, models like K-Nearest Neighbors struggled with scalability and data sparsity, while ensemble methods such as Boosting Regression performed well but required more computational resources. Regularized Linear Regression's balance of precision and efficiency explains its superior results across all metrics. This performance highlights the importance of selecting appropriate models based on data complexity and the trade-offs between accuracy and computational cost.

Feature Importance

Table 3 show that, Study hours emerged as the most significant predictor of GPA, overshadowing other behavioral factors. The feature importance analysis shows that Study_Hours_Per_Day accounted for the largest mean dropout loss (0.367), indicating its substantial contribution to prediction accuracy. Secondary factors, including Extracurricular_Hours_Per_Day (0.203) and Sleep_Hours_Per_Day (0.203) had moderate influence, while Social_Hours_Per_Day (0.202) and Physical_Activity_Hours_Per_Day (0.202) showed minimal impact. These results suggest that academic focused activities are more crucial to GPA outcomes than leisure or physical factors. Understanding these contributions enables targeted strategies to optimize time management and improve academic performance.

Table 3. Feature Importance Metrics

Feature Importance Metrics	Mean dropout loss
Study_Hours_Per_Day	0.367

Extracurricular_Hours_Per_Day	0.203
Sleep_Hours_Per_Day	0.203
Social_Hours_Per_Day	0.202
Physical_Activity_Hours_Per_Day	0.202

The study's dataset size and lack of psychological and environmental variables impose certain limitations on its findings. In particular, the dataset may not adequately represent various student demographics with only 2,000 entries, thereby lowering generalizability. Moreover, stress, mental health, and learning environments were excluded, which potentially eliminates important GPA determinants. It is worth noting that these variables may increase the model's predictive power and applicability. To address these gaps, future studies should increase the dataset and add more factors to enhance predictive accuracy.

8. Conclusion

Regularized Linear Regression emerges as the best predictive model for GPA using behavioral and time-related characteristics within machine learning frameworks. Among the seven models examined, Regularized Linear Regression showed the highest accuracy with a low MSE (0.040) and high R^2 (0.536). This model was particularly effective for this dataset because of its simplicity and precision-interpretability balance. Supporting this conclusion, studies show that time allocation variables, especially study hours, significantly affect GPA. Furthermore, this study highlights that machine learning in education can enable data-driven academic initiatives.

In practical terms, results apply to customized academic support and time management optimization. Specifically, these findings can help educators and advisors design study plans that balance extracurricular and physical activity. By identifying key GPA-related factors, institutions can tailor support to improve student outcomes and lessen academic problems. Ultimately, these consequences demonstrate predictive models' utility in educational planning and pave the way for academic system deployments.

Looking forward, expanding the dataset and adding variables will improve model accuracy and generalizability in future investigations. By including stress, health, and environmental factors, academic performance can be better understood. In addition, a broader and more diversified dataset will increase prediction robustness and guarantee findings apply to diverse student populations. Altogether, these improvements will enhance the prediction framework for academic and institutional use.

Reference

Ali, A., Majeed, M., Saba, K., Bodenarain, A., & Bukhari, M. H. (2013). Effects of Different Sleeping Patterns on Academic Performance in Medical School Students. *Natural Science*, *05*(11), 1193–1198. https://doi.org/10.4236/ns.2013.511146

Al-Khani, A. M., Sarhandi, M. I., Zaghloul, M., Ewid, M., & Saquib, N. (2019). A Cross-Sectional Survey on Sleep Quality, Mental Health, and Academic Performance Among Medical Students in Saudi Arabia. *BMC Research Notes*, *12*(1). https://doi.org/10.1186/s13104-019-4713-2

Ananda, R., & Prasetiadi, A. (2021). Classification Based on Configuration Objects by Using Procrustes Analysis. *Jurnal Infotel*, *13*(2), 76–83. https://doi.org/10.20895/infotel.v13i2.637

Chen, F., & Cui, Y. (2020). Utilizing Student Time Series Behaviour in Learning Management Systems for Early Prediction of Course Performance. *Journal of Learning Analytics*, *7*(2), 1–17. https://doi.org/10.18608/jla.2020.72.1

Creswell, J. D., Tumminia, M. J., Price, S., Sefidgar, Y., Cohen, S., Ren, Y., Brown, J., Dey, A. K., Dutcher, J. M., Villalba, D., Mankoff, J., Xu, X., Creswell, K., Doryab, A., Mattingly, S., Striegel, A., Hachen, D., Martinez, G., & Lovett, M. C. (2023). Nightly sleep duration predicts grade point average in the first year of college. *Proceedings of the National Academy of Sciences*, *120*(8), e2209123120. https://doi.org/10.1073/pnas.2209123120

do Nascimento, R. L. S., das Neves Junior, R. B., de Almeida Neto, M. A., & de Araújo Fagundes, R. A. (2018). Educational Data Mining: An Application of Regressors in Predicting School Dropout. In P. Perner (Ed.), *Machine Learning and Data Mining in Pattern Recognition* (pp. 246–257). Springer International Publishing. https://doi.org/10.1007/978-3-319-96133-0_19

Drzewiecki, W. (2016). Improving Sub-Pixel Imperviousness Change Prediction by Ensembling Heterogeneous Non-Linear Regression Models. *Geodesy and Cartography*, *65*(2), 193–218. https://doi.org/10.1515/geocart-2016-0016

Gao, K., & Li, W. (2019). Prediction and Evaluation of the Severity of Acute Respiratory Distress Syndrome Following Severe Acute Pancreatitis Using an Artificial Neural Network Algorithm Model. *HPB*, *21*(7), 891–897. https://doi.org/10.1016/j.hpb.2018.11.009

Guang, W., Ma, X.-Y., Cheng, X., Luo, C., Wang, H., Xu, X., Lee, K. K., & Yang, X. (2021). The Effects of Long-term Extracurricular Scientific Research on the Medical Students: Insight From Jinan University Medical School. *Biochemistry and Molecular Biology Education*, *49*(4), 535–545. https://doi.org/10.1002/bmb.21499

Guarín, C. E. L., Guzmán, E. L., & González, F. A. (2015). A Model to Predict Low Academic Performance at a Specific Enrollment Using Data Mining. *Ieee Revista Iberoamericana De Tecnologias Del Aprendizaje*, *10*(3), 119–125. https://doi.org/10.1109/rita.2015.2452632

Hakami, A. R. (2021). Effect of Absenteeism on the Performance of Medical Sciences Students: Gender Differences. *Medical Education Online*, *26*(1). https://doi.org/10.1080/10872981.2021.1875531

Hershner, S. (2020). Sleep and academic performance: Measuring the impact of sleep. *Current Opinion in Behavioral Sciences*, *33*, 51–56. https://doi.org/10.1016/j.cobeha.2019.11.009

Hu, J., & Cheung, C. K. J. (2021). Gender Difference in the Effect of Cultural Distance on Academic Performance Among Cross-Border Students in China. *Psicologia Reflexão E Crítica*, *34*(1). https://doi.org/10.1186/s41155-021-00199-4

Iqbal, Z. (2017). *Machine Learning Based Student Grade Prediction: A Case Study*. https://doi.org/10.48550/arxiv.1708.08744

Krishna. (2021). Student's Performance Prediction. *International Journal for Research in Applied Science and Engineering Technology*, *9*(5), 524–528. https://doi.org/10.22214/ijraset.2021.34222

Maia, 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. https://doi.org/10.3390/world4020019

Mak, K.-K., Lee, S.-L., Ho, S.-Y., Lo, W.-S., & Lam, T.-H. (2012). Sleep and Academic Performance in Hong Kong Adolescents. *Journal of School Health*, *82*(11), 522–527. https://doi.org/10.1111/j.1746-1561.2012.00732.x

Negru-Subtirica, O., Pop, E. I., Crocetti, E., & Meeus, W. (2019). Social Comparison at School: Can GPA and Personality Mutually Influence Each Other Across Time? *Journal of Personality*, *88*(3), 555–567. https://doi.org/10.1111/jopy.12510

Nelyub, V. (2023). Machine Learning to Identify Key Success Indicators. *E3s Web of Conferences*, *431*, 05014. https://doi.org/10.1051/e3sconf/202343105014

Pandian, B. R. (2024). Exploring the Role of Machine Learning in Forecasting Student Performance in Education: An in-Depth Review of Literature. *Multidisciplinary Reviews*, *6*, 2023ss043. https://doi.org/10.31893/multirev.2023ss043

Perotta, B., Arantes-Costa, F. M., Enns, S. C., Figueiro-Filho, E. A., Paro, H., Santos, I. S., Lorenzi-Filho, G., Martins, M. A., & Tempski, P. Z. (2021). Sleepiness, sleep deprivation, quality of life, mental symptoms and perception of academic environment in medical students. *BMC Medical Education*, *21*(1), 111. https://doi.org/10.1186/s12909-021-02544-8

Putra, Utami, D. K. I., & Dewa Ayu Agus Sri Laksemi. (2023). Relationship Between Sleep Quality and Learning Achievement in Junior High School Students in Denpasar. *International Journal of Research and Review*, *10*(1), 11–18. https://doi.org/10.52403/ijrr.20230102

Setiyorini, T., & Asmono, R. T. (2020). Implementation of Gain Ratio and K-Nearest Neighbor for Classification of Student Performance. *Jurnal Pilar Nusa Mandiri*, *16*(1), 19–24. https://doi.org/10.33480/pilar.v16i1.813

Shadid, A., Shadid, A. M., Shadid, A., Almutairi, F. E., Almotairi, K. E., Aldarwish, T., Alzamil, O., Alkholaiwi, F., & Khan, S.-U.-D. (2020). Stress, Burnout, and Associated Risk Factors in Medical Students. *Cureus*. https://doi.org/10.7759/cureus.6633

Shahiri, A. M., Husain, W., & Rashid, N. A. (2015). A Review on Predicting Student's Performance Using Data Mining Techniques. *Procedia Computer Science*, *72*, 414–422. https://doi.org/10.1016/j.procs.2015.12.157

Singh, S. P., Malik, S., & Singh, P. (2023). Factors Affecting the Academic Performance of Students. *Jep*. https://doi.org/10.7176/jep/14-15-04

Thorat, Prof. M. (2023). Student Performance Prediction. *International Journal for Research in Applied Science and Engineering Technology*, *11*(10), 1789–1792. https://doi.org/10.22214/ijraset.2023.56313

Tjahyaningtijas, H. P. A. (2023). Machine Learning on Academic Education: Bibliometric Studies. *E3s Web of Conferences*, *450*, 02010. https://doi.org/10.1051/e3sconf/202345002010

Urrila, A. S., Artiges, E., Massicotte, J., Miranda, R., Vulser, H., Bézivin-Frere, P., Lapidaire, W., Lemaître, H., Penttilä, J., Conrod, P. J., Garavan, H., Paillère Martinot, M.-L., & Martinot, J.-L. (2017). Sleep habits, academic performance, and the adolescent brain structure. *Scientific Reports*, *7*(1), 41678. https://doi.org/10.1038/srep41678

Ursache, A., Robbins, R., Chung, A., Dawson-McClure, S., Kamboukos, D., Calzada, E. J., Jean-Louis, G., & Brotman, L. M. (2021). Sleep, Classroom Behavior, and Achievement Among Children of Color in Historically Disinvested Neighborhoods. *Child Development*, *92*(5), 1932–1950. https://doi.org/10.1111/cdev.13590

Yağcı, M. (2022). Educational data mining: Prediction of students' academic performance using machine learning algorithms. *Smart Learning Environments*, *9*(1), 11. https://doi.org/10.1186/s40561-022-00192-z

Yang, H., & Xiang, J. (2024). Peer support and academic engagement: The moderated mediation model for college students. *Psychology in the Schools*, pits.23324. https://doi.org/10.1002/pits.23324

Zeng, H. Z. (2023). Influence of Physical Activity, SES`, Perceived Safety, and Demographic Factors on the GPA of Asian American Students: Report of the National High School Longitudinal Study. *International Journal of Research in Education and Science*, *9*(3), 623–637. https://doi.org/10.46328/ijres.3155

Zúñiga-Prado, J. R., Rodríguez-Paz, M. L., Herrera-Rubio, A. M., & Osorio, S. (2023). Influence of Sociodemographic Factors on Academic Performance in the Subject of Human Gross Anatomy. *International Journal of Morphology*, *41*(1), 96–103. https://doi.org/10.4067/s0717-95022023000100096