A Data-Driven Predictive Analysis of Academic Performance and Study Efficiency Among College Students

Romer M. Real[1]; Jay T. Castor[2]; Rolly B. Villanueva[3]; Reginaldo Morijon[4]; Cesar Joey Santillan[5]; Serafin C. Palmares[6]; Kristine T. Soberano[7]

[1,2,3,4,5,6,7]State University of Northern Negros, Philippines

Publication Date: 2026/05/16

Abstract: This study examined the academic performance and study efficiency of college students, focusing on identifying critical behavioral predictors of success. The research was initially planned as a primary descriptive-survey using a 30-item Likert-scale questionnaire (Cronbach's α = .85) for students in Negros Occidental. Due to ethical considerations during examination week, the study utilized a secondary data-driven dataset, a sample of 150 students was analyzed using a descriptive-predictive design with descriptive statistics and linear regression. Findings revealed an average exam score of 70.59%, with study hours identified as the strongest predictor (B = 0.870), followed by attendance (B = 0.499) and motivation level (B = 0.686), all showing positive effects on academic performance. This study makes a number of contributions that was not previously being studied on an academic level using the Study Efficiency metric, defined as output per unit of time. The analysis determined a Study Efficiency ratio of 4.07 on average, emphasizing that an efficient behavioral strategy makes more difference to academic achievement than the time available for study.

Keywords: *Academic Performance; Study Habits; Time Management; Predictive Analytics; Study Efficiency*.

How to Cite: Romer M. Real; Jay T. Castor; Rolly B. Villanueva; Reginaldo Morijon; Cesar Joey Santillan; Serafin C. Palmares; Kristine T. Soberano (2026) A Data-Driven Predictive Analysis of Academic Performance and Study Efficiency Among College Students. *International Journal of Innovative Science and Research Technology*, 11(5), 344-348. https://doi.org/10.38124/ijisrt/26may654

I. INTRODUCTION

In the modern academic community, identifying the reasons behind student success has become a critical challenge for educational institutions. Academic performance is no longer viewed as a product of intelligence alone but as a complex interplay of behavioral habits, environmental factors, and time management strategies. As college students face increasing academic pressure, the need for a "Data-Driven Predictive Analysis Approach" becomes essential to move beyond traditional education methods and toward objective, actionable insights.

Consequently, this study utilized data-driven approach with curated dataset of 150 students to simulate and test predictive models. By shifting to secondary data, the research applied rigorous statistical techniques, such as linear regression, to determine "Study Efficiency" defined as the ratio of academic output (exam scores) relative to time input (study hours). This approach allows for a more agile analysis of how variables like attendance and resource availability contribute to an 85% productivity threshold, which was identified as a benchmark for student success in preliminary observations.

By analyzing these data-driven predictors, this study aimed to provide a framework that students and educators can use to refine educational habits instead of simply increasing study duration, the findings offer a predictive model that emphasizes the optimization of specific behaviors to maximize academic efficiency without compromising student well-being during intense academic periods. Furthermore, this study positions "Study Efficiency" as the central innovation and primary contribution of the research. Unlike traditional approaches that focus only on study duration, this metric quantifies academic output relative to time investment, offering a practical and measurable way to evaluate how effectively students convert effort into performance. By introducing this metric, the study advances educational predictive analytics by shifting the focus from quantity of study to quality and efficiency of learning.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26may654

II. OBJECTIVES

The primary goal of this research is to develop and evaluate a data-driven predictive model to identify the key determinants of academic performance and study efficiency among college students. Specifically, the study aims to achieve the following:

- To analyze the demographic and behavioral profile of students in terms of age, attendance percentage, and weekly study hours.
- To determine the level of academic performance of the 150-student sample based on exam scores and final grades.
- To establish a quantitative measure for "Study Efficiency" by calculating the ratio of academic output (Exam Scores) relative to time investment (Study Hours).
- To identify significant predictors of success among variables such as attendance, motivation levels, and resource availability using linear regression analysis.
- To evaluate the impact of environmental factors, such as internet access and the use of educational technology, on the 85% productivity benchmark.
- To propose a predictive framework that allows students to optimize their study habits and maximize efficiency without increasing academic stress.

III. LITERATURE REVIEW

Additional studies have explored alternative predictive frameworks, including decision-support systems and efficiency-based models, which further support the role of data-driven approaches in academic forecasting (Lagman, 2015; Weston, 2020; Ji et al., 2024). Recent literature suggests that academic performance is influenced by a synergy of internal motivation and external support systems (Abdallah & Abdullah, 2020; Ampomah et al., 2024).

A. Predictors of Academic Performance

Research by global educational theorists highlights that attendance and time management remain the strongest predictors of student outcomes. According to contemporary studies, students with high attendance rates consistently demonstrate a more profound understanding of course materials, leading to higher exam scores. Furthermore, the availability of resources, including internet access and educational technology (EduTech), has become a prerequisite for success in the digital learning era. Studies indicate that "digital readiness" significantly buffers the impact of socioeconomic disparities on student grades (Feng et al., 2025).

B. The Concept of Study Efficiency

A critical gap in traditional literature is the distinction between study duration and study quality. While many studies focus on the "number of hours studied," few provide a measurable way to evaluate how effectively time is converted into academic performance. This study addresses this gap by introducing the Study Efficiency metric, which captures the relationship between time investment and academic output. Recent research further supports this concept by emphasizing the importance of measuring actual learning gains per unit of time rather than relying solely on study duration. It is observed that students who maintain an 85% productivity threshold often rely on focused, high-intensity study sessions rather than prolonged, low-concentration efforts. This study adopts this efficiency-centric view to evaluate how students optimize their limited study time during high-pressure periods such as examination weeks.

C. Data-Driven Predictive Analytics in Education

The shift toward predictive modeling in education reflects a broader trend in "Educational Data Mining." By utilizing regression analysis and machine learning algorithms, researchers can forecast student outcomes with high accuracy (Al-Obaidi, 2023; Luo et al., 2024; Zuo et al., 2025). The transition from primary data collection to curated secondary datasets is a recognized practice in large-scale academic research. This methodology allows for the analysis of high-granularity data without the logistical delays and ethical risks of distracting students during critical academic milestones. Using such datasets provides a robust baseline for developing predictive frameworks that can be applied to localized settings, such as universities in Negros Occidental.

IV. MATERIALS AND METHODS

This section outlines the research design, participants, and the technical procedures used to analyze the factors influencing academic performance and study efficiency.

A. Research Design

This study employed a descriptive-predictive research design. It combines descriptive statistics to profile student behaviors and inferential statistics (linear regression) to establish predictive relationships. The design was structured to evaluate "Study Efficiency," a derived metric quantifying the relationship between time investment and academic outcomes.

B. Initial Research Plan and Methodological Pivot

The study was initially designed as a primary survey using a validated 30-item Likert-scale instrument (Cronbach's $\alpha = .85$). However, due to its overlap with the students' examination period, the study adopted a secondary dataset to ensure ethical compliance and avoid disruption. This streamlined approach enabled a focused application of predictive modeling while maintaining data reliability.

C. Participants and Sampling

A sample of 150 students was extracted from the curated dataset through a randomized selection process to ensure a representative distribution of behavioral variables. The sample includes data on student age, attendance rates, study hours, and academic scores (Exam Scores and Final Grades).

D. Data Analysis and Predictive Modeling

The collected data were processed using the following statistical methods:

- Descriptive Statistics: Frequency, mean, and standard deviation were used to profile student attendance, study hours, and exam performance.
- Metric Derivation: "Study Efficiency" was calculated for each participant using the formula: Study Efficiency =

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26may654

Exam Score ÷ Study Hours. This ratio represents the number of percentage points achieved on the examination per hour of weekly study time.

- Predictive Analysis: A Linear Regression model was deployed to identify the weight of individual predictors (e.g., Attendance and Motivation) in forecasting the Final Grade.
- Software: Data processing and visualization were conducted using Python, utilizing standard data analysis libraries for statistical computation and regression modeling.

E. Ethical Considerations

The pivot to a secondary dataset was primarily an ethical decision to avoid interfering with students' mental well-being and concentration during their examination week. All data used were anonymized to ensure the privacy and confidentiality of the student records, adhering to international data privacy standards for educational research.

V. RESULTS AND DISCUSSION

This section presents the findings of the predictive analysis conducted on the 150-student sample, focusing on the profile of academic performance and the efficiency of study habits.

A. Student Profile and Academic Performance

The descriptive analysis revealed that the average Exam Score for the sample was 70.59%, with scores ranging from a minimum of 41% to a maximum of 100%. The mean Attendance rate was 79.27%, while students spent an average of 19.35 hours per week on dedicated study. Table 1 provides the complete descriptive profile of the participants.

Table 1. Descriptive Statistics of Student Performance and Behavior (N = 150)

Variable	Mean	Std. Deviation	Min	Max
Age	23.83	3.49	18.00	29.00
Attendance (%)	79.27	11.78	60.00	99.00
Study Hours (Weekly)	19.35	6.23	8.00	38.00
Exam Score (%)	70.59	16.80	41.00	100.00
Study Efficiency Ratio	4.07	1.71	1.45	10.44

B. Analysis of Correlation and Study Efficiency

A key contribution of this study is the introduction of the "Study Efficiency" metric, which quantifies academic output per hour of study. This metric provides a practical way to evaluate how effectively students convert study time into academic performance. The analysis established a mean Study Efficiency ratio of 4.07. This indicates that, on average, for every hour of weekly study, students gained approximately 4.07 percentage points in their examination scores. As shown in Figure 1, while there is a positive trend between hours and scores, the high variability implies that quality of study is a more potent predictor than time volume alone.

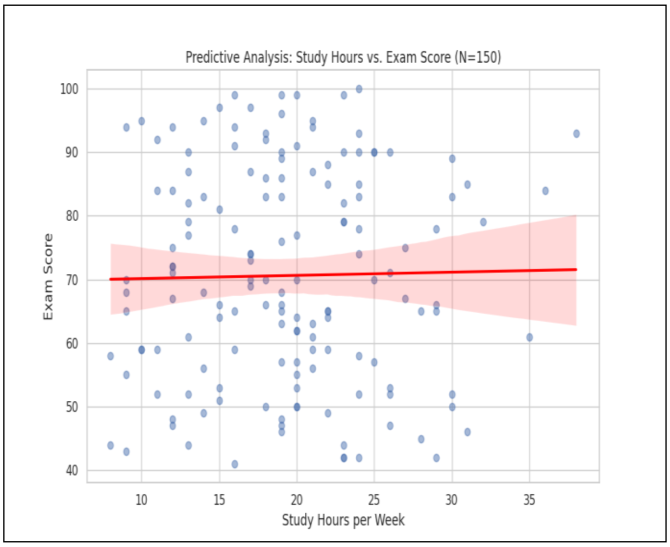

Fig 1. Predictive Analysis: Relationship between Weekly Study Hours and Exam Scores

Figure 1 shows a positive but variable relationship between study hours and exam scores, indicating that increased study time does not always guarantee higher performance, suggesting that study efficiency plays a more critical role than study duration alone in predicting academic outcomes.

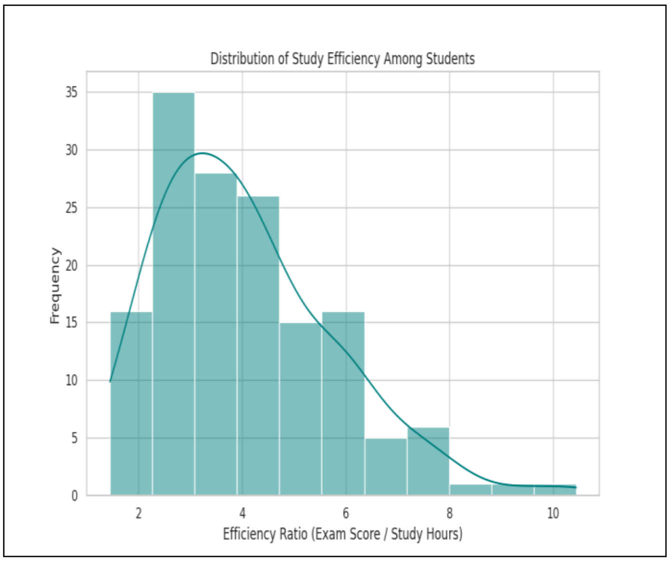

Fig 2. Distribution of Study Efficiency Among Students

Figure 2 illustrates the distribution of study efficiency, showing that most students fall within a moderate efficiency range, suggesting that only a subset of students achieve high-efficiency learning, which may explain variations in academic performance despite similar study hours.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26may654

D. *Predictive Modeling of Academic Success*

Linear regression was employed to identify the weight of various predictors on the Final Grade. The model revealed that Study Hours, Attendance, and Motivation Level were significant positive predictors. Table 2 summarizes the impact of each variable within the predictive model.

Table 2. Linear Regression Coefficients for Predicting Academic Performance

Predictor Variable	Coefficient (B)	Impact Direction
(Constant)	10.85	Baseline Score
Attendance (%)	0.499	Positive
Study Hours	0.870	Positive
Motivation Level	0.686	Positive

Interestingly, the model also highlighted that while motivation levels vary, their impact on performance is mediated by the consistency of assignment completion and resource availability.

The regression model demonstrated strong explanatory power ($R^2 = 0.769$; Adjusted $R^2 = 0.764$), indicating that approximately 76.9% of the variance in academic performance is explained by the selected variables. The overall model was statistically significant ($F = 161.7$, $p < 0.001$), confirming that the predictors reliably explain variations in exam scores.

Among the variables, study hours ($B = 0.870$, $p < 0.001$) emerged as the strongest predictor, followed by attendance ($B = 0.499$, $p < 0.001$). Motivation level also showed a statistically significant positive effect ($B = 0.686$, $p = 0.015$). The regression model can be expressed as follows: Exam Score = 10.85 + 0.499(Attendance) + 0.870(Study Hours) + 0.686(Motivation).

The results further indicated that students maintaining the 85% productivity benchmark often possessed stable internet access and high engagement with educational technology (EduTech). The strong influence of attendance may be attributed to increased exposure to structured instruction, direct interaction with educators, and real-time clarification of concepts, which enhance learning effectiveness compared to independent study alone. This finding is consistent with previous studies that identify attendance and time management as key predictors of academic success (Abdallah & Abdullah, 2020; Ampomah et al., 2024). Additionally, recent research in educational data analytics highlights that structured study behavior and time allocation are strong determinants of performance outcomes, particularly in predictive modeling contexts (Luo et al., 2024; Zuo et al., 2025).

E. *Discussion of Findings*

The findings are consistent with the “Problem Engineer” perspective of optimizing systems to achieve maximum output. The results show that behavioral factors such as study hours and attendance significantly influence academic performance, which aligns with behaviorist and cognitive learning theories that emphasize repetition, reinforcement, and structured engagement. Increased study time strengthens memory through repeated exposure, while consistent attendance provides guided instruction and immediate feedback that support deeper understanding. However, the dominance of study hours and attendance as predictors can be explained through behavioral reinforcement mechanisms and structured learning exposure, and their effectiveness depends on the quality of engagement. This indicates that time-on-task alone is insufficient without cognitive efficiency, reinforcing the role of Study Efficiency as a mediating construct that captures both behavioral input and actual learning effectiveness. These findings are consistent with previous studies highlighting the importance of time management and engagement in academic success (Abdallah & Abdullah, 2020; Ampomah et al., 2024; Luo et al., 2024), while also extending the discussion by showing that academic success is shaped not only by effort, but by how learning activities are structured and executed.

Unlike prior predictive analytics studies that emphasize model accuracy, this study integrates behavioral interpretation, demonstrating not only which variables predict performance but also how efficiently these variables translate into measurable outcomes. Furthermore, the strong predictive capability of the model highlights the potential of data-driven approaches in developing early-warning systems that can identify at-risk students using indicators such as declining attendance, reduced study hours, and low Study Efficiency ratios. These indicators can guide targeted interventions, including personalized tutoring, attendance monitoring, and structured study programs, enabling institutions to intervene before performance declines become critical. The observed variability in Study Efficiency ratios further supports its value as a practical metric for distinguishing differences in learning effectiveness, suggesting that educational strategies should focus on “smarter study” through improved time management and resource optimization rather than simply increasing study duration. By leveraging these objective measures, academic institutions in Negros Occidental can implement proactive monitoring systems that identify students with declining engagement or low efficiency before the final examination period.

VI. CONCLUSION AND RECOMMENDATION

A. *Conclusion*

This study successfully developed a data-driven predictive model to evaluate academic performance and study efficiency among college students. By analyzing a sample of 150 students, the research established that while the average exam score stands at 70.59%, academic success is heavily influenced by high-impact variables such as study hours and attendance, highlighting that both time investment and consistent class participation are critical drivers of student performance. The introduction of the Study Efficiency metric represents the primary innovation of this research, providing a dual contribution by advancing theoretical understanding of learning efficiency while offering a practical, quantifiable measure for assessing academic productivity and distinguishing between high-effort and high-impact learning strategies.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26may654

Furthermore, the methodological pivot from primary survey collection to secondary data analysis proved to be a critical ethical and technical success. It allowed the researchers to generate high-granularity insights while respecting the students' academic focus during their examination week. The findings suggest that a "Data-Driven Predictive Analysis Approach" is a robust framework for identifying early indicators of academic risk and providing objective benchmarks for student productivity. This study contributes to the field of educational predictive analytics by providing a practical framework for identifying at-risk students and supporting data-driven academic interventions.

This positions the study as an Educational Predictive Analytics Model for Early Intervention, providing institutions with a practical framework for improving student outcomes through data-driven decision-making.

B. Recommendation

Based on the findings of this study, the following recommendations are proposed:

- For Students: Shift focus from increasing study hours to improving "Study Efficiency." Students should prioritize high-engagement behaviors, such as maintaining high attendance and utilizing educational technology, to maximize their 85% productivity threshold.
- Educators: Implement data-driven early warning systems that monitor attendance and engagement metrics. This allows for proactive interventions before high-stakes testing periods like exam weeks.
- School Administrators: Invest in digital infrastructure, such as stable internet and EduTech tools, as these were identified as critical predictors of academic success in the predictive model.
- Future Researchers: Conduct longitudinal studies using the validated 30-item Likert-scale questionnaire (Cronbach's $\alpha = .85$) during non-examination periods to compare localized primary perceptions with the secondary data-driven findings established in this study.

REFERENCES

[1]. Abdallah, N., & Abdullah, A. (2020). Predicting student performance using data mining and learning analytics techniques: A systematic literature review. Applied Sciences, 11(1), 237. https://doi.org/10.3390/app11010237.

[2]. Al-Obaidi, H. (2023). Student grade prediction using machine learning methods. Rochester Institute of Technology. https://repository.rit.edu/cgi/viewcontent.cgi?article=13370&context=these.

[3]. Ampomah, O., Agyekum, L., Akuoko-Frimpong, J., & Quansah, A. (2024). Predicting students' academic performance via machine learning algorithms: An empirical review and practical application. Computer Engineering and Intelligent Systems, 15(1). https://doi.org/10.7176/CEIS/15-1-09.

[4]. Feng, J., Yu, B., Tan, W. H., Dai, Z., & Li, Z. (2025). Key factors influencing educational technology adoption in higher education: A systematic review. Retrieved from https://pmc.ncbi.nlm.nih.gov/articles/PMC12040101.

[5]. Ji, M., Le, J., Chen, B., & Li, Z. (2024). A predictive model for classifying college students' academic performance based on visual-spatial skills. Frontiers in Psychology, 15, 1434015. https://doi.org/10.3389/fpsyg.2024.1434015.

[6]. Lagman, A. (2015). Predictive decision support system using logistic regression and decision tree model combination for student graduation success determination. https://doi.org/10.21016/IRRC.2015.AU05EF82O.

[7]. Luo, Z., Mai, J., Feng, C., Kong, D., Liu, J., Ding, J., Qi, B., & Zhu, Z. (2024). A method for prediction and analysis of student performance that combines multi-dimensional features of time and space. Mathematics, 12(22), 3597. https://doi.org/10.3390/math12223597.

[8]. Weston, T. (2020). Efficiency in higher education: A quantitative analysis of academic effort and grade outcomes. Journal of Learning Analytics, 7(2), 45–58.

[9]. Zuo, M., Wang, K., Tang, P., Meng, X., & Luo, H. (2025). Predicting academic performance from future-oriented daily time management behavior: A LASSO-based study of first-year college students. Behavioral Sciences, 15(9), 1242. https://doi.org/10.3390/bs15091242.