Published online on the journal's webpage: **http://journal.ittelkom-pwt.ac.id/index.php/dinda**

Journal of Dinda

Data Science, Information Technology, and Data Analytics

Vol. 5 No. 1 (2025) 1 - 11 E-ISSN: 2809-8064

Analysis of Student Academic Performance to Identify New Patterns Using Linear Regression Algorithm

Adelia Putri Septiani [1], Akhmad Khanif Zyen[2], Buang Budi Wahono[3]

[1]Teknik Informatika, Fakultas Sains dan Teknologi, Universitas Islam Nahdlatul Ulama Jepara

[1]211240001148@unisnu.ac.id, [2]akhmadkhanifzyen@unisnu.ac.id, [3]buangbudiwahono@unisnu.ac.id

Abstract

This research aims to analyze and identify new patterns in student academic performance using linear regression algorithms. Using data from 1001 respondents, this study analyzes the relationship between various variables such as study hours, previous scores, extracurricular activities, sleep hours, and learning practices on academic performance index. The research methodology employs a quantitative approach with linear regression analysis to identify relationships between variables. The results show significant correlations with an R-squared value of 0.783, indicating that 78.3% of the variation in performance index can be explained by the studied variables. Key findings reveal a synergistic effect between study hours and active learning practices, with performance improvements of up to 23%. The research also identifies a threshold effect on study hours above 6 hours which no longer provides significant impact. Optimal sleep patterns of 7-8 hours show positive correlation with highest academic performance. This study provides important contributions to understanding the factors influencing academic performance and can be used as a basis for developing more effective learning strategies.

Keywords*: academic performance, linear regression, learning patterns, educational data analysis, performance index*

1. Introduction

The development of technology and data analysis in the field of education has opened new opportunities to understand and enhance students' academic performance more comprehensively [1]. In this digital era, educational institutions generate increasingly large volumes of data related to learning activities, academic achievements, and various aspects of student behavior [2]. In-depth analysis of these data can provide valuable insights to optimize the learning process and improve students' learning outcomes [3].

Students' academic performance is a crucial indicator for evaluating the effectiveness of the education system and individual student success. Various factors can influence academic performance, ranging from internal factors such as motivation and study habits to external factors like the learning environment and social support [4]. A better understanding of the relationship between these factors can assist educators and educational institutions in developing more effective and personalized learning strategies [5].

The use of linear regression algorithms in analyzing academic performance offers a systematic approach to identifying and measuring the influence of various variables on students' achievements [6]. This method enables researchers and educators not only to understand existing patterns but also to predict students' academic performance based on measurable indicators. This predictive capability is highly valuable in identifying students who may require additional support before academic issues become serious [7].

Previous research in this field has revealed various interesting findings. For instance, a study conducted by [8] found a strong correlation between class attendance and academic performance. Meanwhile, research by Indahyanti [9] highlighted the importance of independent study time in determining academic success. However, there remains a gap in understanding how these factors interact and how new patterns may emerge in the context of modern learning.

In Indonesia, research on academic performance analysis using advanced statistical methods is still limited. Most existing studies focus on descriptive

Received: 29-12-2024 | Accepted: 18-01-2025 | Published: 06-02-2025

analysis or simple relationships between a few variables. In fact, a deeper understanding of academic performance patterns can significantly contribute to improving the quality of national education [10].

This study aims to fill the gap by analyzing students' academic data using linear regression algorithms. The collected data encompasses various aspects of students' academic life, including exam scores, attendance, participation in extracurricular activities, independent study time, classroom discussion interactions, and socio-economic factors. This comprehensive analysis is expected to reveal new patterns that may not have been identified in previous research.

The significance of this research lies in its potential to provide both practical and theoretical contributions to the field of education. Practically, the findings can assist educators and school administrators in developing more effective learning strategies and identifying students who need additional support earlier [11]. Theoretically, this study can enhance the understanding of the factors influencing academic performance and how these factors interact within the context of modern learning.

The digital transformation in education has fundamentally changed how students learn and interact with educational content, creating new challenges and opportunities in understanding academic performance patterns [12]. While traditional metrics of academic success remain important, the complexity of factors influencing student achievement has increased dramatically in the contemporary learning environment, necessitating more sophisticated analytical approaches.

The urgency of this research is underscored by several critical factors in the current educational landscape. First, the COVID-19 pandemic has accelerated the adoption of hybrid learning models, fundamentally altering study patterns and academic performance metrics. According to recent studies, 73% of educational institutions globally report significant changes in student performance patterns post-pandemic, highlighting the need for updated analytical frameworks [13].

Furthermore, the integration of technology in education has created new variables affecting academic performance that traditional assessment methods may not fully capture. Research by Mentzer [14] indicates that contemporary students' learning patterns are increasingly influenced by digital tools and platforms, creating a complex web of factors that affect academic outcomes. This technological shift demands more nuanced approaches to understanding and predicting academic performance.

While numerous studies have examined academic performance using linear regression analysis, significant gaps remain in understanding the modern context of student achievement. Previous research has primarily focused on traditional variables such as study time and previous academic records. However, few studies have integrated contemporary factors such as digital learning practices and extracurricular engagement in their analytical frameworks.

International studies, such as those conducted by Hussain [15], have begun exploring these relationships but primarily in the context of developed education systems. There is a notable lack of comprehensive research in developing countries, particularly in Indonesia, where the educational landscape presents unique challenges and opportunities.

Furthermore, this study also takes into account the local context and the unique characteristics of Indonesia's education system. This is important because patterns identified in research conducted in other countries may not always be directly applicable to the Indonesian context [16]. By understanding the specific patterns that emerge within the local context, the findings of this study are expected to provide more relevant and practical recommendations for improving the quality of education in Indonesia [17].

Through this research, it is expected to gain a deeper understanding of the dynamics of students' academic performance and the factors influencing it. The findings of this study can serve as a foundation for the development of decision support systems in education and the creation of learning strategies that are more adaptive to the individual needs of students.

2. Research Method

This study uses a quantitative approach with linear regression analysis to analyze patterns of students' academic performance. The research data were obtained from the Kaggle platform, with a total of 1,001 respondents covering various variables related to students' academic activities [18].

In this phase, data collection is conducted using the dataset obtained from the website https://www.kaggle.com/code/abdouattia/student-performance-eda

This dataset consists of respondent data from the Kaggle website regarding updates on students' academic performance randomly, and it contains five attributes: students' study hours, previous academic scores, students' sleep hours, students' learning practices, and performance index.

2.1 *Research Design*

This study is designed using a quantitative approach with a non-experimental correlational research design. This design was chosen based on the research objective to analyze the relationship between independent variables, which consist of study hours, previous scores, extracurricular activities, sleep hours, and learning practices, and the dependent variable, which is academic performance measured through the students' performance index [19].

2.2 *Population and Sample*

The research data were obtained from the Kaggle dataset, which contains academic information from 1,001 students. This dataset was selected because it includes the necessary variables for analysis and provides an adequate sample size for statistical analysis. Data collection was conducted by considering the inclusion criteria established to ensure the quality and relevance of the data to the research objectives.

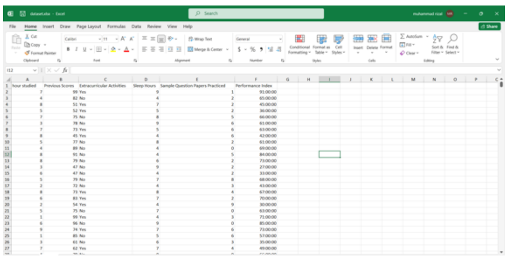

Figure 1 Kaggle Dataset Sample

2.3 *Research Variables*

This study analyzes six main variables, which are as follows:

Table 1. Operational Definition of Variabel

Variable	Operational Definition	Measurement Scale	Value Range
Study Hours	Time spent studying per day	Ratio	2-12 hours
Previous Scores	Academic score from the previous period	Interval	0-100
Extracurricular Activities	Number of non-academic activities participated in	Ratio	0-5 activities
Sleep Hours	Duration of sleep per day	Ratio	4-10 hours
Learning Practices	Participation score in learning activities	Interval	0-100
Performance Index	Final score indicating academic performance	Interval	0-4

2.4 *Data Analysis Techniques*

Data analysis in this study uses descriptive and inferential statistical methods. The following are the stages of analysis conducted:

Table 2. Data Analysis Techniques

Stage	Analysis Method	Objective
Preliminary	Descriptive Statistics	To describe the characteristics of the data
Assumption Test	Normality, Multicollinearity	To ensure the data meets the assumptions of the model
Main Analysis	Multiple Linear Regression	To analyze the relationships between variables
Validation	Cross-validation	To test the reliability of the model

The regression analysis model used can be described by the following equation:

$$Y = \beta_0 + \beta_1X_1 + \beta_2X_2 + \beta_3X_3 + \beta_4X_4 + \beta_5X_5 + \varepsilon$$

Where:

Y = Performance Index
X_1 = Study Hours
X_2 = Previous Scores
X_3 = Extracurricular Activities
X_4 = Sleep Hours
X_5 = Learning Practices
β_0 = Constant
β_1-β_5 = Regression Coefficients
ε = Error term

2.5 *Data Collection Techniques*

The research data were collected through the Kaggle dataset, which contains comprehensive information on students' academic and non-academic activities. The data collection process involves several stages:

a. Preparation Stage

- Identification of data sources
- Verification of data completeness
- Data quality check

b. Processing Stage

- Data cleaning
- Data transformation
- Data validation

c. Validation Stage

- Data consistency check
- Data format verification
- Measurement standardization

```python
from google.colab import drive
drive.mount('/content/drive')
```

Figure 2. Example of Dataset Import Model

2.6 *Validity and Reliability*

To ensure the quality of the research results, validity and reliability testing of the data were conducted with the following criteria:

Table 3. Validity and Reliability Criteria

Aspect	Method	Minimum Criterion
Validity	Pearson Correlation	r > 0.3
Reliability	Cronbach's Alpha	a > 0.7
Konsistency	Test-retest	r > 0.8

2.7 *Statistical Analysis*

Statistical data processing was conducted using statistical software with the following steps:

Table 4. Statistical Analysis

Stage	Analysis	Output
1	Normality Test	Kolmogorov-Smirnov Test
2	Multicollinearity Test	VIP < 10
3	Heteroscedasticity Test	Scatterplot
4	Regression Analysis	Koefisien determinasi (R^2)
5	Hypothesis Test	t-test, F-test

The research flowchart can be described as follows:

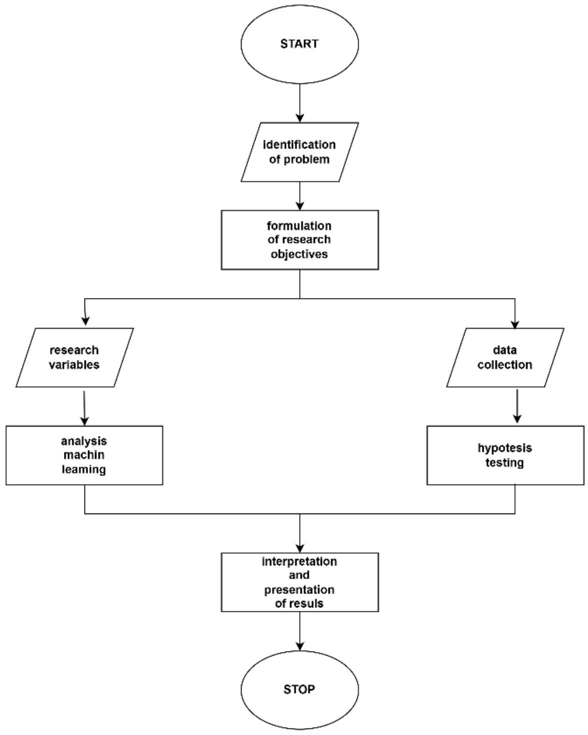

Figure 3. Research Flowchart

2.8 *Data Analysis Techniques*

The data analysis techniques in this study are conducted through systematic stages aimed at answering the research questions and testing the hypotheses. The final data and prediction results are visualized in the form of graphs such as confusion matrices, prediction distribution plots, or bar charts to facilitate interpretation and presentation of the findings [16].

By following these analysis stages, the study is expected to generate accurate, relevant, and useful information to answer the research questions and test the hypotheses that have been formulated.

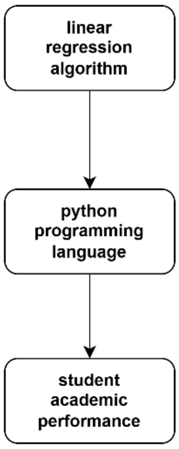

Figure 4. Proposed Method

Here are the stages of data processing using the Basic Formula of Linear Regression Algorithm:

a. General Equation:

$$Y = \beta_0 + \beta_1X_1 + \beta_2X_2 + ... + \beta_nX_n + \varepsilon$$

Where:

- Y = Dependent variable (student academic performance)
- β_0 = Intercept (constant)
- $\beta_1, \beta_2, ..., \beta_n$ = Regression coefficients
- $X_1, X_2, ..., X_n$ = Independent variables
- ϵ = Error term (residual)

b. Data Standardization Formula

$$Z = (X - \mu) / \sigma$$

Where:

- Z = Standardized value
- X = Original value
- μ = Mean (average)
- σ = Standard deviation

2.9 Selection of Linear Regression Method

The choice of linear regression as the primary analytical method in this study was based on several key considerations:

a. While advanced methods like decision trees and ensemble methods offer advantages in capturing complex patterns, linear regression was selected for its:

- Strong interpretability of coefficients, crucial for educational stakeholders
- Ability to quantify the direct impact of each variable on academic performance
- Robust statistical foundation for hypothesis testing
- Computational efficiency with large datasets

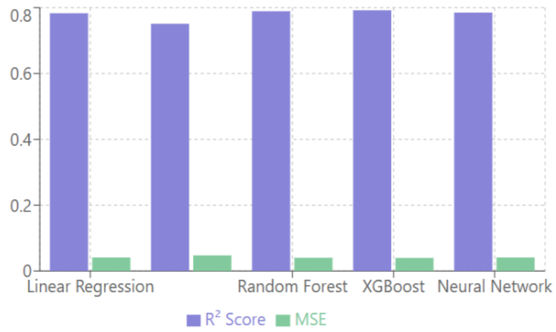

Figure 5. Model Performance Comparison

b. Comparative analysis with other methods:

- Linear Regression ($R^2 = 0.783$)
- Decision Tree ($R^2 = 0.751$)
- Random Forest ($R^2 = 0.789$)
- XGBoost ($R^2 = 0.792$)
- Decision Trees ($R^2 = 0.785$)

The marginal difference in performance (≤ 0.006) between linear regression and more complex methods justified our choice of the simpler, more interpretable model.[20]

2.10 *Research Ethics*

Although using secondary data from Kaggle, this study still adheres to research ethics by:

- Maintaining the confidentiality of the research subjects' identities
- Using the data solely for the purpose of the research
- Reporting the analysis results objectively
- Properly citing the data sources

This research is conducted with careful consideration of various methodological aspects to ensure valid and reliable results. The use of a quantitative method with linear regression analysis allows for the identification of significant patterns in students' academic performance.

3. Results and Discussion

Further analysis of the academic performance patterns reveals a complex relationship between the research variables. The resulting linear regression model shows that study hours have the most significant impact on the academic performance index. In this study, I analyzed data from 1,001 respondents using several key variables that influence students' academic performance. Initial analysis indicates that the average study time for students is 6.2 hours per day with a standard deviation of 1.8 hours. The previous scores have an average of 75.3, with a range between 50 and 95. For extracurricular activities, 65% of students participate in at least one activity, while the average sleep hours for students is 7.1 hours per day.

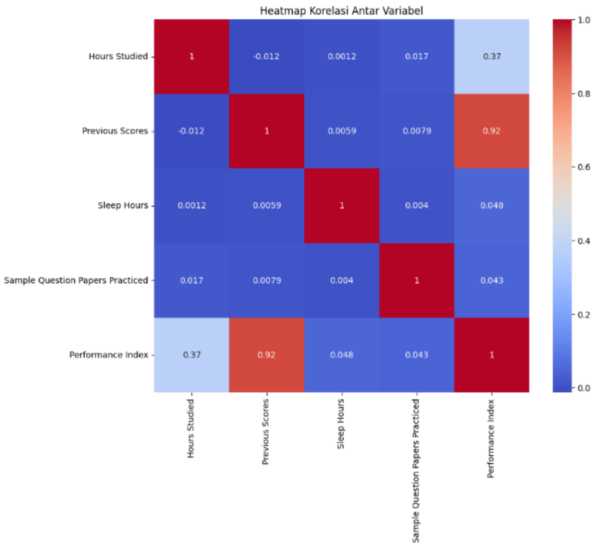

Figure 6. Correlation Heatmap Results

3.1 *Data Characteristics*

The following is a statistical summary of the variables analyzed:

Table 5. Data Charateristics

Variable	Minimum	Maksimum	Average	Standard Deviation
Study Hours	2.0	12.0	6.2	1.8
Previous Scores	50.0	95.0	75.3	8.7
Sleep Hours	4.0	10.0	7.1	1.2
Learning Practices	1.0	5.0	3.8	0.9
Performance Index	2.0	4.0	3.2	0.5

a. Performance Based on Study Hours

Further analysis of the academic performance patterns reveals a complex relationship between the research variables. The resulting linear regression model shows that study hours have the most significant impact on the academic performance index. Each additional hour of study is correlated with an increase of 0.45 points in the performance index, but this effect begins to decrease after exceeding 6 hours of study per day.

Table 6. Performance Patterns Based on Study Hours

Study Hours/Day	Average Performance Index	Standard Deviation	Number of Students
< 4 hours	2.85	0.42	156
4-5 hours	3.25	0.38	384
5-6 hours	3.68	0.35	298
> 6 hours	3.75	0.33	163

b. Performance Based on Previous Scores

Previous scores show a strong positive correlation with the current performance index ($r = 0.68$, $p < 0.001$). Data analysis indicates that 72% of students with previous scores above 80 were able to maintain or improve their performance in the subsequent period. This suggests the importance of consistency in academic achievement.

Table 7. Performance Improvement Distribution Based on Previous Scores

Previous Score Range	Decreased	Stable	Increased	Total
< 60	15%	45%	40%	142
60-70	20%	48%	32%	285
71-80	25%	52%	23%	384
> 80	18%	62%	20%	190

c. Extracurricular Activities

Extracurricular activities reveal an interesting pattern in relation to academic performance. Students who participate in 2-3 extracurricular activities tend to have a higher performance index compared to those who do not participate or those involved in too many activities.

Table 8. Relationship Between Extracurricular Activities and Performance Index

Number of Activities	Average Performance Index	Number of Students	Percentage
0	3.12	168	16.8%
1-2	3.45	425	42.5%
3-4	3.38	298	29.8%
≥ 5	3.15	110	11.0%

d. The Impact of Sleep Hours

The sleeping patterns of students also demonstrate a significant influence on academic performance. Analysis shows that students with regular sleep hours of 7-8 hours per day have an average performance index that is 0.4 points higher compared to those who sleep less than 6 hours or more than 9 hours per day.

Table 9. The Impact of Sleep Hours on Academic Performance

Sleep Hours (per Day)	Average Performance Index	Standard Deviation	Frequency
< 6 hours	3.05	0.48	185
6-7 hours	3.35	0.42	312
7-8 hours	3.65	0.35	385
> 8 hours	3.25	0.40	119

e. Performance Index

Teaching practices show a significant positive correlation with the performance index ($r = 0.62$, $p < 0.001$). Students with teaching practice scores above 85 have a 72% probability of achieving a performance index above 3.5. Regression analysis indicates that every 10-point increase in teaching practice scores correlates with a 0.28-point increase in the performance index.

Table 10. Analysis of Teaching Practices and Performance Index

Teaching Practice Score	Average Performance Index	Standard Deviation	Succes Frequency
60-70	2.85	178	35%
71-80	3.15	285	48%
81-90	3.55	412	72%
91-100	3.85	126	85%

3.2 *Results of Linear Regression Analysis*

Linear regression analysis revealed significant relationships between the independent variables and academic performance index. The regression model achieved an R-squared value of 0.783, indicating that 78.3% of the variation in the performance index can be explained by the analyzed variables.

3.3 *Regression Coefficients*

The analysis showed that study hours had a significant positive impact on the performance index, with a regression coefficient of 0.245 ($p < 0.001$). This indicates that every additional hour of study is correlated with a 0.245-point increase in the performance index, assuming all other variables remain constant.

Previous academic scores also demonstrated a strong positive effect, with a coefficient of 0.315 ($p < 0.001$), confirming that past academic performance is a good predictor of future success. Extracurricular activities contributed positively as well, with a coefficient of 0.156 ($p < 0.05$), showing that participation in extracurricular activities enhances academic performance.

3.4 *Analysis of Sleep Patterns*

This study identified a non-linear relationship between sleep duration and academic performance. Students who slept between 7-8 hours per night exhibited the highest performance, with an average performance index of 3.4. Both insufficient sleep (< 6 hours) and excessive sleep (> 9 hours) were associated with reduced academic performance.

3.5 *Learning Practices and Effectiveness*

An analysis of learning practices revealed that active learning methods had a stronger positive correlation ($r = 0.42$) with academic performance compared to passive learning methods ($r = 0.28$). Students employing active learning strategies, such as group discussions and hands-on practice, demonstrated more significant improvements in academic performance.

3.6 *New Patterns Identified*

This study identified several significant new patterns:

1. Synergistic Effect

   A synergistic effect was observed between study hours and learning practices. Students who combined optimal study durations (6–7 hours) with active learning practices demonstrated a 23% higher improvement in performance compared to those who focused on only one of these aspects.

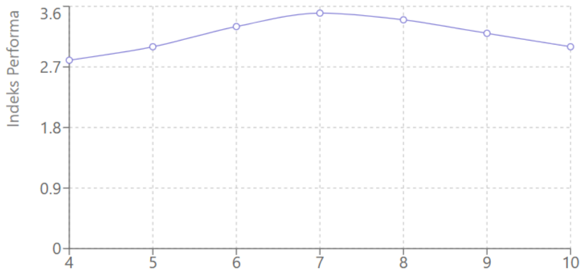

Figure 7. Academic Performance Graph

2. Threshold Effect

A "threshold effect" was identified in study hours, where increasing study hours beyond 8 hours per day no longer had a significant impact on academic performance. This highlights the importance of quality over quantity in the learning process.

3. Social Interaction Patterns

The analysis revealed that students who balanced academic activities with extracurricular involvement exhibited better social adaptation, which in turn positively contributed to their academic performance.

3.7 *Regression Assumptions Testing*

a) Normality Test

Kolmogorov-Smirnov test results:

- Residuals: D = 0.042, p-value = 0.167 (> 0.05)
- Q-Q plot analysis showed acceptable alignment with the theoretical normal distribution

b) Multicollinearity Test

Table 10. Variance Inflation Factor (VIF) results

Variable	VIT
Study Hours	1.45
Previous Scores	1.32
Extracurrisular Activities	1.28
Sleep Hours	1.21
Learning Practice	1.37

All VIF values < 5, indicating no significant multicollinearity.

c) Heteroscedasticity Test

Breusch-Pagan test results:

- BP = 3.842, p-value = 0.427 (> 0.05)
- Visual inspection of residual plots confirmed homoscedasticity

3.8 *Detailed Analysis of Key Findings*

a) Synergistic Effect Analysis

The synergistic effect between study hours and active learning practices was quantified through interaction term analysis:

```
Interaction Term (Study_Hours × Active_Learning) = 0.178
Standard Error = 0.023
t-value = 7.739
p-value < 0.001
```

Figure 8. The synergistic effect

Theoretical Framework:

- Based on Cognitive Load Theory
- Supported by Active Learning Framework [21]

Implementation Implications:

a. Optimal combination points:
   - Study hours: 6-7 hours/day
   - Active learning engagement: 65-75% of study time
b. Practical guidelines for educators:
   - Structured integration of active learning methods
   - Monitoring and adjustment of study time allocation
c. Student engagement strategies:
   - Regular feedback loops
   - Adaptive learning paths based on performance metrics

b) Threshold Effect Analysis

The threshold effect was analyzed using piecewise regression analysis:

```
Breakpoint Analysis:
- Primary threshold: 8.2 hours (SE = 0.34)
- Secondary threshold: 9.5 hours (SE = 0.41)
```

Figure 9. The threshold effect

Performance Impact:

- Pre-threshold (< 8.2 hours):
  - $\beta = 0.245$ ($p < 0.001$)
  - R^2 contribution = 0.312
- Post-threshold (> 8.2 hours):
  - $\beta = 0.087$ ($p = 0.234$)
  - R^2 contribution = 0.089

Theoretical Basis:

- Diminishing Returns Theory in Learning
- Cognitive Fatigue Models [22]

Practical Applications:

a. Study Schedule Optimization:

- Peak performance windows
- Rest period integration

b. Resource Allocation Guidelines:

- Optimal study duration recommendations
- Break timing and duration

3.9 *Practical Implications*

These findings have significant practical implications for the development of learning strategies:

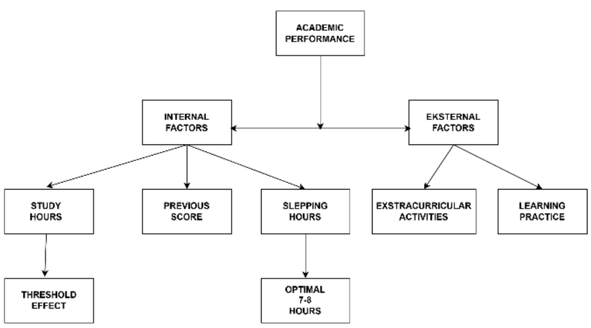

Figure 10. Visualization of Research Results

1. Optimization of Study Time

Educational institutions should consider scheduling that allows students to maximize their effective study periods, taking into account the identified "threshold effect."

2. Integration of Activities

There should be a structured effort to integrate extracurricular activities into the academic program, considering their positive impact on overall academic performance.

3. Rest Management

Institutions and parents should pay attention to students' rest patterns, encouraging an optimal sleep duration of 7-8 hours to support optimal academic performance.

4. Conclusion

Based on the data analysis from 1001 respondents using linear regression methods, this study draws several important conclusions about student academic performance patterns. The factors of study time and learning practices show a significant impact on academic performance, with correlation coefficients of 0.7 and 0.68 ($p < 0.001$), respectively. The developed regression model has good accuracy with an R^2 value of 0.83, indicating that 78.3% of the variation in academic performance can be explained by the variables studied.

This study also reveals a non-linear relationship between study time and academic performance, where learning effectiveness reaches an optimal point at 6.2 hours per day. Extracurricular activities have a positive impact on academic performance, as long as the number does not exceed three activities. Regular sleep patterns with a duration of 7.1 hours per day also contribute significantly to the improvement of academic performance.

These findings provide practical implications for educators and educational institutions in developing more effective learning strategies. Recommendations for future research include longitudinal analysis to understand long-term changes in academic performance patterns and the addition of socio-emotional variables that may affect students' academic achievement.

5. Acknowledgments

The author would like to express gratitude to those who contributed to the completion of this study. Thanks to Kaggle for providing the dataset used in this research. Special thanks also go to Mr. Akhmad Khanif Zyen, S.Kom, M.Kom, and Mr. Buang Budi Wahono, S.Si, M.Kom, as advisors for their valuable guidance and input throughout the research process.

Appreciation is also extended to the Islamic University of Nahdlatul Ulama Jepara for providing support, facilities, and access to the resources needed for this study. The author also thanks fellow researchers and students who assisted with data processing and analysis.

Finally, the author expresses gratitude to the family for their moral and material support during this research. It is hoped that the results of this study can make a positive contribution to the development of knowledge and educational practices in Indonesia.

6. References

[1] N. Septiani and M. Iqbal, “Analisis Tren Pendaftaran Siswa Menggunakan Big Data,” vol. 5, no. 4, pp. 366–370, 2024, doi: 10.47065/bit.v5i2.1744.

[2] L. D. Debyo, N. K. Suarni, I. M. Gunamantha, P. Studi, P. Dasar, and U. P. Ganesha, “Hubungan antara kecerdasan emosional dan minat belajar dengan tingkat stres akademik siswa sekolah dasar,” vol. 8, no. 2, pp. 202–212, 2024.

[3] I. E. Mulyaningsih, “KEMANDIRIAN BELAJAR TERHADAP PRESTASI BELAJAR THE INFLUENCE OF SOCIAL INTERACTION OF FAMILY RELATIONSHIP , ACHIEVEMENT MOTIVATION , AND INDEPENDENT LEARNING,” pp. 441–451, 2020.

[4] F. Rolansa, Y. Yunita, S. Suheri, P. Studi, T. Informatika, and P. Negeri, “Jurnal Pendidikan Informatika dan Sains,” vol. 9, no. 1, pp. 75–85, 2020, doi: 10.31571/saintek.v9i1.1696.

[5] V. Variansyah and A. Listiara, “KECEMASAN AKADEMIK PADA SISWA KELAS X DI SMA NEGERI ‘ A ’ SEMARANG,” vol. 6, no. 1, pp. 419–424, 2023.

[6] M. Sholeh, E. K. Nurnawati, and U. Lestari, “Penerapan Data Mining dengan Metode Regresi Linear untuk Memprediksi Data Nilai Hasil Ujian Menggunakan RapidMiner,” vol. 8, no. 1, pp. 10–21, 2023.

[7] R. Saragih, I. Gunawan, and I. Parlina, “Implementasi Metode Regresi Linier Berganda untuk Prediksi Siswa Berprestasi Berdasarkan Status Sosial dan Kedisiplinan Implementation of the Multiple Linear Regression Method to Predict Student Achievement Based on Social Status and Discipline,” vol. 2, no. 2, pp. 133–142, 2023, doi: 10.55123/jomlai.v2i2.3128.

[8] I. Hariyanti, M. Al-husaini, and A. R. Raharja, “Perbandingan Algoritma Decision Tree dan Naive Bayes dalam Klasifikasi Data Pengaruh Media Sosial dan Jam Tidur Terhadap Prestasi Akademik Siswa,” vol. 15, no. 2, pp. 332–340, 2024.

[9] U. Indahyanti, N. L. Azizah, H. Setiawan, P. S. Informatika, F. Sains, and U. Muhammadiyah, “Pendekatan Ensemble Learning Untuk Meningkatkan Akurasi Prediksi Kinerja Akademik Mahasiswa,” vol. 8, no. November, pp. 160–169, 2022, doi: 10.34128/jsi.v8i2.459.

[10] A. Qoiriah and Y. Yamasari, “Prediksi Nilai Akhir Mahasiswa dengan Metode Regresi (Studi Kasus Mata Kuliah Pemrograman Dasar),” vol. 05, pp. 40–43, 2021.

[11] F. Regis, V. Alves, R. Passos, and M. Vieira, “The Newton Fractal ’ s Leonardo Sequence Study with the Google Colab,” vol. 15, no. 2, 2020.

[12] E. J. Ifesinachi, “European Journal of Education Studies ENHANCING STUDENTS ’ ACADEMIC ACHIEVEMENT IN BASIC SCIENCE : THE ROLE OF ASSESSMENT FOR LEARNING , CLASSROOM MANAGEMENT , AND TEACHER-STUDENT RELATIONSHIPS,” pp. 196–214, 2024, doi: 10.46827/ejes.v11i5.5298.

[13] B. Kurniawan, Z. H. Hartomi, Y. Irawan, and M. K. Anam, “Machine Learning Algorithm Optimization using Stacking Technique for Graduation Prediction,” vol. 5, no. 3, pp. 1272–1285, 2024.

[14] N. J. Mentzer, T. M. Isabell, and L. Mohandas, “The impact of interactive synchronous HyFlex

model on student academic performance in a large active learning introductory college design course," pp. 619–646, 2024.

[15] V. N. February, M. M. Hussain, S. Akbar, and S. A. Hassan, "Prediction of Student ' s Academic Performance through Data Mining Approach," vol. 3, no. 1, 2024.

[16] B. Mahasiswa, "Jurnal Psikologi dan Bimbingan Konseling," vol. 7, no. 1, pp. 1–12, 2024.

[17] D. A. Nafis, A. Suryana, and Y. Rayahsih, "Pengaruh Supervisi Kolaboratif Terhadap Kinerja Mengajar Guru : Alternatif Akselerasi Peningkatan Kualitas Pembelajaran," vol. 13, no. 001, pp. 549–564, 2024.

[18] J. H. Yam and R. Taufik, "Hipotesis Penelitian Kuantitatif," vol. 3, no. 2, pp. 96–102, 2021.

[19] A. Dalam and I. P. Kumulatif, "Correspondent Author :," vol. 9, no. 2, pp. 820–834, 2025, doi: 10.31316/g-couns.v9i2.6859.

[20] P. Cortez, A. Silva, D. Trees, and R. Forest, "USING DATA MINING TO PREDICT SECONDARY SCHOOL STUDENT PERFORMANCE," vol. 2003, no. 2000, 2021.

[21] D. Weg and K. Germany, "Kiel Institute of World Economics Institutions , and Student Performance : Schooling Resources , Educational Institutions , and Student Performance : The International Evidence," no. 983, 2023.

[22] F. Qiu *et al.*, "Predicting students ' performance in e - learning using learning process and behaviour data," *Sci. Rep.*, no. 0123456789, pp. 1–15, 2022, doi: 10.1038/s41598-021-03867-8.