ISSN: 2658-8455
Volume 6, Issue 12 (2025), pp. 575-594.

Utilizing learning analytics to analyze and predict student academic performance

Abdelkhalek ZINE
National School of Commerce and Management
University Mohammed First of Oujda

Abdelali KAAOUACHI
National School of Commerce and Management
University Mohammed First of Oujda

Correspondence address :	National School of Commerce and Management University Mohammed First of Oujda Complexe universitaire BP 658, Oujda 60000 05365-06983
Disclosure Statement :	The authors declare that they have not received any financial support that could have influenced the objectivity of this study. They take full responsibility for any potential plagiarism, the use of artificial intelligence in the writing process, as well as for the results presented in this article.
Conflict of Interest :	The authors report no conflicts of interest.
Cite this article :	ZINE, A., & KAAOUACHI, A. (2025). Utilizing learning analytics to analyze and predict student academic performance. International Journal of Accounting, Finance, Auditing, Management and Economics, 6(12), 575–594. https://doi.org/10.5281/zenodo.17768015
License	**This is an open access article under the CC BY-NC-ND license**

International Journal of Accounting, Finance, Auditing, Management and Economics - IJAFAME
ISSN: 2658-8455
Volume 6, Issue 12 (2025)

Utilizing learning analytics to analyze and predict student academic performance

Abstract

This study seeks to evaluate students' academic performance across various subjects, identify disparities in educational outcomes, and explore the individual, familial, institutional, and socio-economic factors that influence these performances. By doing so, it aims to provide a comprehensive understanding of the dynamics affecting student achievement. Data were gathered from the ELMOSSALA school within the provincial directorate of Meknes for the year 2024. The analysis involved examining exam results to evaluate students' mastery of fundamental skills in each subject area. Statistical tests were utilized to compare performance across different subjects. Additionally, multiple regression analysis was conducted to identify influential factors, and advanced data analysis techniques were employed to uncover predictive characteristics linked to academic success. Various machine learning algorithms were also assessed for their efficacy in predicting student performance. The results revealed notable variations in student performance across subject areas, with significant disparities observed between them. Key factors impacting academic performance included socio-economic status, the quality of teaching, and individual student characteristics. Furthermore, some machine learning algorithms demonstrated greater effectiveness than others in predicting learners' performance, highlighting the potential of these technologies in educational settings. While this study provides valuable insights, it is important to acknowledge certain limitations. The research is based on data from a single school, which may limit the generalizability of the findings to broader educational contexts. Additionally, the study focuses primarily on quantitative measures, potentially overlooking qualitative aspects that could further enrich the understanding of student performance.

Keywords: academic performance, educational disparities, influencing factors, fundamental skills, regression analysis, data analysis, machine learning.
Classification JEL : C80
Paper type: Empirical Research

ISSN: 2658-8455
Volume 6, Issue 12 (2025), pp. 575-594.

1. Introduction

Education is widely recognized as a key driver of human development and socio-economic progress, providing individuals with the foundational competencies necessary to participate effectively in contemporary knowledge-based societies. However, international assessments such as PISA and PIRLS consistently reveal persistent learning gaps in many countries, particularly in the mastery of fundamental skills in reading, mathematics, and scientific literacy (OECD, 2019). Research also shows that learning deficits at the primary level can have long-term consequences on academic trajectories, employability, and social mobility (UNESCO, 2020), reinforcing structural inequalities.

In Morocco, several studies have highlighted persistent challenges in basic education despite ongoing reforms. National evaluations conducted by the Ministry of National Education, as well as international analyses, show that a significant proportion of learners struggle to acquire essential competencies in languages, mathematics, and scientific reasoning (CSEFRS, 2018). Similar findings were reported by Sobhi, Sophie, and Amapola (2010), who note that the Moroccan education system continues to face major obstacles related to skill acquisition and learning quality. These difficulties threaten not only students' academic progression but also the effectiveness and equity of the educational system.

Given these challenges, the integration of learning analytics and educational data mining has emerged as a promising avenue for improving decision-making and identifying at-risk learners. Studies such as Arnold and Pistilli (2012) and Romero and Ventura (2013) demonstrate the capacity of predictive models to detect students likely to encounter academic difficulties, enabling early and targeted interventions. Other research, including that of Rizvi et al. (2019), emphasizes the influence of demographic and socio-economic factors on student outcomes, advocating for analytical approaches that take into account both cognitive and contextual variables.

Within this context, ELMOSSALA School seeks to adopt a data-driven approach to better understand the determinants of student performance and to strengthen pedagogical practices. Traditional descriptive analyses, though useful, remain insufficient to capture the complexity of the factors influencing academic success. This study thus proposes a comprehensive analytical framework combining descriptive statistics, inferential modeling, and machine learning techniques to examine academic performance during the first semester of 2024 for a sample of 1,292 students across six grade levels.

The present article is structured to address the central research question—What factors influence students' academic performance, and how can predictive models support early identification of at-risk learners?—through a sequence of interconnected analyses. First, a literature review synthesizes existing research on learning analytics, determinants of academic success, and predictive modeling in education. Second, the methodology details the data collection process, variables considered, and analytical techniques employed, including multiple regression and machine learning algorithms. Third, the results present an in-depth examination of performance disparities across subjects and the influence of socio-demographic factors. Finally, the discussion critically interprets the findings in light of the literature and proposes implications for pedagogical practice and institutional decision-making.

By grounding the introduction in established research and clearly outlining the analytical pathway, this study contributes to both the institutional context of ELMOSSALA School and the broader scientific literature on educational analytics. It offers actionable insights for improving student performance while reinforcing the methodological rigor essential for scientific publication.

2. Literature Review

The body of literature on academic performance reveals a wide spectrum of methodological approaches and analytical perspectives, providing a robust foundation for understanding the mechanisms underlying student achievement. Early studies such as those by Arnold and Pistilli (2012) and Burgos et al. (2018) converge on the importance of early prediction and timely intervention for students at risk of academic failure. However, their methodological orientations differ markedly: Arnold and Pistilli focus primarily on digital trace data extracted from learning management systems, whereas Burgos et al. incorporate socio-behavioral indicators, yielding models with greater explanatory depth but reduced external validity across diverse institutional contexts. Likewise, the work of Rizvi et al. (2019) highlights the role of socio-economic and geographical characteristics in shaping academic outcomes, echoing insights from broader international reports such as UNESCO (2020). Yet, these studies rely on decision-tree models that contrast with the more systemic and large-scale frameworks often used in global educational analyses, underscoring the challenge of reconciling micro-level and macro-level determinants.

Research on virtual learning environments further demonstrates considerable methodological diversity. While Romero and Ventura (2013) use educational data mining techniques to show that interactivity and learner engagement positively influence performance, Huang and Fang (2013) adopt quasi-experimental designs based on interactive e-books to reach similar conclusions. The consistency of their findings highlights the importance of learner engagement, but the divergence in methodological choices illustrates the limitations of generalizing results across educational levels and instructional modalities.

Recent work in educational data mining has placed increasing emphasis on machine learning techniques. Studies by Okubo et al. (2017), Jha et al. (2019), and Hassan et al. (2019) demonstrate that algorithms such as logistic regression, artificial neural networks, and other supervised models can effectively identify at-risk students and support the design of personalized interventions. Although these approaches successfully uncover complex and nonlinear patterns in large datasets, concerns remain regarding the transparency and interpretability of such models, particularly in educational contexts where stakeholders require clear justifications for data-driven decisions. Complementary research comparing modeling approaches—including logistic and linear regression (Barber & Sharkey, 2012; Roberge et al., 2012), neural networks (Calvo-Flores et al., 2006), Bayesian networks (Pardos et al., 2007), decision trees (Baker & Yacef, 2009), and deep learning architectures (Urrutia & Cartwright, 2017; Wang et al., 2017)—illustrates the trade-off between predictive accuracy and interpretability. More recent contributions, such as Karimi et al. (2020), propose knowledge graph–based relational models to capture the multifaceted interactions among cognitive, socio-economic, and institutional variables, highlighting the trend toward integrated and holistic frameworks.

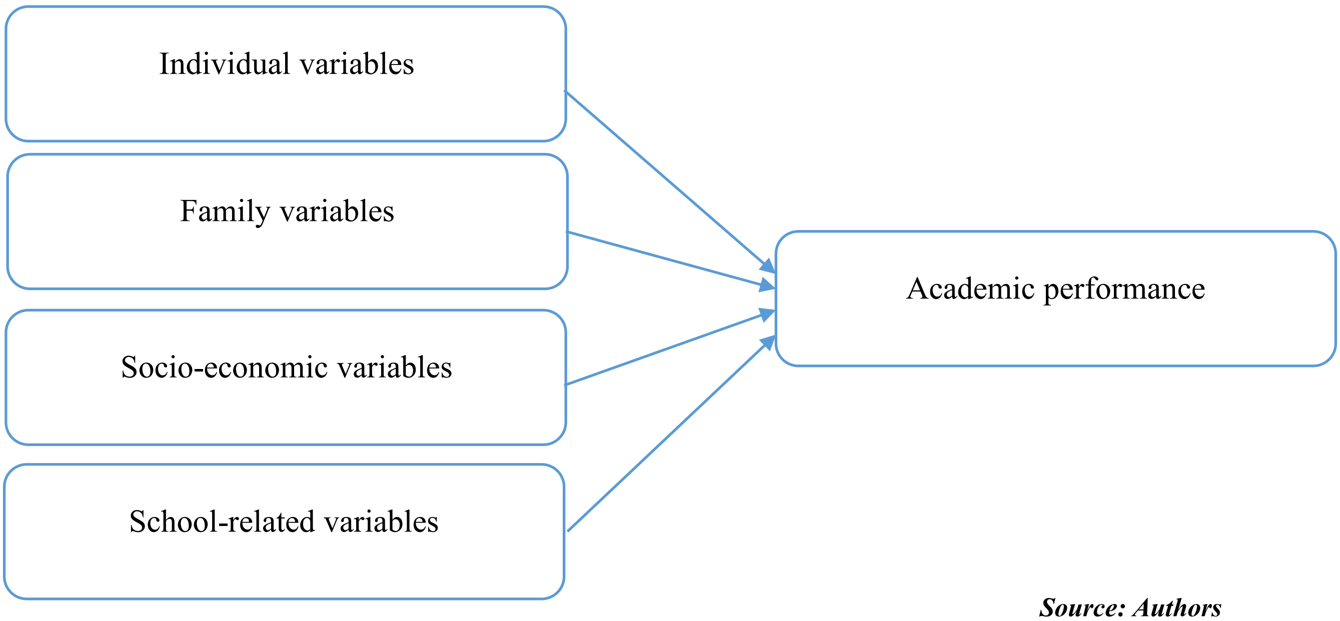
Taken together, this literature demonstrates the growing centrality of data-driven approaches in understanding and predicting academic performance. Nevertheless, several limitations persist, particularly the insufficient attention paid to institutional context—such as teacher quality, school climate, resource availability, and local policy environments—which may substantially shape student outcomes. Moreover, the challenge of developing models that are simultaneously accurate, interpretable, and contextually adaptable remains unresolved.

Building on these theoretical foundations, the present study formulates several hypotheses aimed at empirically examining the complex interplay of individual, familial, socio-economic, and academic variables. First, previous research suggests that student achievement varies significantly across subject areas due to differences in cognitive demands and instructional practices; thus, we hypothesize that academic performance will differ across disciplines (H1). Second, demographic and familial characteristics—including gender, parental education, and

ISSN: 2658-8455
Volume 6, Issue 12 (2025), pp. 575-594.

family structure—are expected to significantly influence academic outcomes (H2), consistent with findings from Rizvi et al. (2019) and UNESCO (2020). Third, socio-economic conditions are anticipated to explain a substantial share of the variance in students' academic results (H3). Fourth, mastery of fundamental disciplinary skills should directly contribute to overall academic performance, indicating that stronger achievement in core subjects predicts higher global averages (H4). Finally, following recent developments in learning analytics, we expect machine learning models to effectively predict student success, with logistic regression yielding superior performance compared to Support Vector Machines and K-Nearest Neighbors (H5).
In light of these hypotheses, the theoretical framework guiding this study is organized around four major categories of determinants consistently identified in the literature as key predictors of academic performance. The first comprises individual-level variables, including gender, study habits, and academic difficulties. The second concerns family-related factors, such as parental education and family structure. The third category encompasses socio-economic indicators, including socio-economic status and housing conditions. The fourth focuses on school-related variables, particularly students' mastery levels in different subject areas, which collectively shape their overall academic trajectory. This conceptual structure provides a coherent basis for analyzing the multifactorial determinants of academic achievement and for developing predictive models aligned with both theoretical insights and empirical evidence.

Fig 1 : Conceptuel model



Source: Authors

3. Methodology

The study was conducted at ELMOSSALA School, a primary institution located in a semi-urban environment, using data from the 2024 first-semester academic results extracted from the MASSAR system, including subject-specific scores, continuous assessment grades, and the overall semester average. These academic records were complemented by a social survey collecting demographic information such as age and gender, family characteristics including parental education and family structure, and socio-economic indicators such as housing conditions and access to educational resources at home. The final sample consists of 1291 students, providing sufficient statistical power for inferential analysis. The research is grounded in a conceptual framework assuming that academic performance is influenced by individual, family, socio-economic, and academic variables, all interacting to explain variations in achievement. To examine these relationships, the study employs a multiple regression model formally expressed as

$$Y_i=\beta_0+\beta_1 IND_i+\beta_2 FAM_i+\beta_3 SES_i+\beta_4 ACAD_i+\epsilon_i$$

where Y_i represents the overall academic performance, IND includes variables such as age and gender, FAM includes parental education and family structure, SES refers to socio-economic indicators, and ACAD includes subject scores and competency mastery indicators. The coefficients $\beta_0 ... \beta_4$ are estimated to determine the strength and direction of the effects, and ϵ_i represents the error term. These variables were selected based on extensive evidence showing their relevance in predicting student achievement. Multiple regression is used because it identifies the relative influence of each independent variable on a continuous dependent variable, making it appropriate for modeling academic outcomes. Principal Component Analysis (PCA) is applied when reducing the dimensionality of correlated socio-economic and academic variables to retain the maximum amount of information. Machine learning algorithms such as Random Forests and Support Vector Machines are incorporated not as primary explanatory tools but for predictive comparison, allowing the study to evaluate the robustness of traditional statistical models against advanced predictive approaches. The data were processed using statistical software such as SPSS, R, or Python. Processing steps included cleaning the dataset to treat missing values and outliers, performing descriptive statistics to understand distributions and variances, computing correlation matrices to identify preliminary associations, estimating the regression model, applying PCA, implementing predictive models, and evaluating their accuracy through cross-validation. This methodological approach ensures rigor, coherence, and alignment with the study's objectives.

4. Results

In this section, we present the results of our study on the academic performance of students at ELMOSALA School for the first semester of 2024, as well as the analyses conducted to address our research questions. Each question is approached methodically and in detail, providing clear information on the various analyses performed and the conclusions drawn.

We begin by examining the academic performance of students in each field of study, assessing their mastery of fundamental skills. We present the results of the evaluations conducted and discuss the observed trends.

Next, we analyze the existence of significant differences in the mastery of fundamental skills across the various subjects taught at ELMOSALA School. Appropriate statistical tests are used to compare performance in each discipline.

We also explore demographic, familial, environmental, and institutional factors that are closely correlated with students' academic performance. The results of our correlation analyses are presented and discussed to elucidate their significance.

We identify the most predictive characteristics of academic success among the demographic, familial, environmental, and institutional variables, highlighting the factors most influential on students' academic performance.

Similarly, we evaluate the effectiveness of different machine learning models in predicting the academic performance of students at ELMOSALA School. The results of our modeling analyses are presented, and the performance of various algorithms is compared.

Finally, we compare the different models in terms of accuracy, recall, and F1 score when predicting academic performance. We discuss the strengths and limitations of each model and provide recommendations for their future use.

4.1. Descriptive statistics

The descriptive statistics provide an initial overview of students' demographic characteristics and academic performance, revealing several important patterns within the dataset. On average, students come from moderately sized households, with a mean of two siblings (SD = 0.88),

ISSN: 2658-8455
Volume 6, Issue 12 (2025), pp. 575-594.

IJAFAME
International Journal of Accounting, Finance, Auditing, Management & Economics

which is consistent with common family structures in the Moroccan context. Additional academic support appears limited, as reflected by the low average number of extra hours spent at school (M = 0.41, SD = 0.49), indicating that remedial or reinforcement sessions are not widely utilized. Similarly, home study time remains modest, with students reporting an average of only 0.59 hours of revision per day (SD = 0.88). This low level of extracurricular academic engagement may play a role in shaping students' overall performance.

Table 1: descriptive statistics

Variable	N	Mean	Std	Min	25%	50%	75%	Max
Nombre de freres	1291	2.00	0.88	0	1	2	3	4
Heures supplementaires	1291	0.41	0.49	0	0	0	1	1
Heures de revision a la maison	1265	0.59	0.88	0	0	0	1	3
Arabe	1291	6.31	1.69	0	5.00	6.50	7.67	9.89
Français	1291	6.12	1.49	2.00	5.00	6.06	7.25	9.62
Math	1291	6.25	2.06	1.00	5.00	6.00	8.00	10.00
Éducation islamique	1291	6.84	1.84	0	5.53	7.00	8.00	10.00
Éducation physique	1291	7.21	1.14	0	7.00	7.00	8.00	10.00
Activites scientifiques	1291	6.66	1.81	1.00	5.50	6.75	8.00	10.00
Éducation artistique	1291	6.96	1.27	2.00	6.00	7.00	8.00	10.00
Moyenne generale	1291	6.62	1.35	1.92	5.63	6.63	7.62	9.91

Source: Authors

Regarding academic achievement, students demonstrate varying levels of mastery across subjects. The highest averages are observed in Physical Education (M = 7.21, SD = 1.14) and Islamic Education (M = 6.84, SD = 1.84), subjects that typically require less complex cognitive processing. Performance in core academic subjects is more heterogeneous. In Arabic, the average score reaches 6.31 (SD = 1.69), whereas the mean score in French is slightly lower at 6.12 (SD = 1.49), reflecting persistent challenges in foreign language learning among students in Moroccan public schools. Mathematics shows a mean of 6.25 (SD = 2.06), with a relatively high standard deviation indicating considerable variability in skills and possibly unequal access to academic support.

Intermediary performance levels are observed in transversal subjects such as Scientific Activities (M = 6.66, SD = 1.81) and Artistic Education (M = 6.96, SD = 1.27). The overall grade point average for students is 6.62 (SD = 1.35), with values ranging from 1.92 to 9.91. This wide distribution, coupled with the substantial standard deviation, reveals significant heterogeneity within the student population. While a portion of students achieves satisfactory outcomes, a non-negligible number clearly experiences academic difficulties.

Taken together, these descriptive results suggest a learning environment marked by limited out-of-school academic investment, notable disparities in Mathematics performance, relative strength in non-linguistic subjects, and a general average positioned just above the passing threshold. These findings justify conducting further correlational and multivariate analyses to identify the factors most strongly associated with academic achievement and to better understand the mechanisms underlying students' performance.

4.2. Academic performance of ELMOSALA students

To address the question, "What are the academic performances of ELMOSALA School students in the first semester of 2024, in terms of mastery of fundamental skills in each field of study?", we calculated the percentages of students who achieved the passing grade in each subject. The following tables summarize these academic performances, providing an overview of the results obtained by students across different fields of study.

Table 2: Gender Distribution

	Total	Men	Women
Number	1292	655	637

Source: Authors

The distribution by gender is as follows: out of a total of 1292 individuals, there are 655 males and 637 females.

- **Performance of the entire establishment**

Table 3: performance of the entire establishment

	Arabic Language	French Language	Mathematics	Islamic Education	Physical Education	Scientific Activity	Art Education	Average
Students with more than 5	980	996	1000	1050	1191	1203	1192	1184
Percentage	0,758513932	0,770897833	0,773993808	0,812693498	0,921826625	0,931114551	0,922600619	0,916408669
Not mastered < 5	311	295	291	241	100	88	99	107
Percentage Not mastered	0,240712074	0,228328173	0,225232198	0,186532508	0,077399381	0,068111455	0,076625387	0,082817337
Average 5<6	219	311	178	251	127	86	197	183
Percentage Average	0,169504644	0,240712074	0,137770898	0,194272446	0,098297214	0,066563467	0,15247678	0,141640867
Mastered > 6	761	685	822	799	1064	1117	995	1001
Percentage Mastered	0,589009288	0,530185759	0,63622291	0,618421053	0,823529412	0,864551084	0,770123839	0,774767802

Source: Authors

Table 3 illustrates the academic performance of students across various subjects, providing detailed information on pass rates and levels of skill mastery.

The results show that the subjects with the highest success rates are Scientific Activity (93.11%) and Physical Education (92.18%), closely followed by Artistic Education (92.26%). This indicates a strong overall performance by students in these areas. Conversely, Arabic has the lowest pass rate, with only 75.85% of students scoring above 5, suggesting a need for improvement in this subject.

Arabic also has the highest percentage of students who have not mastered the subject, at 24.07%. It is followed by French (22.83%) and Mathematics (22.52%). On the other hand, the subjects with the lowest non-mastery rates are Scientific Activity (6.81%) and Physical Education (7.74%). This implies that both Arabic and French require pedagogical interventions to reduce the number of struggling students.

A significant group of students falls into the average score range (between 5 and 6). French has the highest percentage in this category at 24.07%, followed by Islamic Education at 19.43%. In contrast, Scientific Activity has the lowest percentage in this range, with only 6.66%. These results indicate that while some students have not fully mastered the subjects, they are in an intermediate zone that could be improved with additional effort.

The subjects where skills are best mastered are Scientific Activity (86.46%) and Physical Education (82.35%), indicating a high level of student proficiency in these areas. Conversely, Arabic has the lowest skill mastery percentage at 58.90%, necessitating increased attention to improve student performance in this subject.

These results highlight the need for specific educational initiatives in subjects where students face the most difficulties while reinforcing the strengths identified in other subjects.

- **Performance of 6th year**

Table 4 provides a detailed overview of students' academic performance across various subjects, highlighting several key aspects of their academic success.

Overall, the results reveal a strong performance by students, with pass rates often exceeding 90% in most subjects. Notably, Physical Education and Artistic Education stand out with a 100% pass rate, indicating excellent skill mastery in these areas.

ISSN: 2658-8455
Volume 6, Issue 12 (2025), pp. 575-594.

Table 4: performance of 6th year

	Arabic language	French language	Social Sciences	Mathematics	Islamic education	Physical Education	Scientific activity	Art education	Average
Average Observed	7,860960699	6,961659389	8,172576419	7,454061135	8,173580786	8,558951965	7,748908297	8,227074236	7,894934498
Students with more than 5	222	198	226	211	228	228	220	229	227
Percentage	0,969432314	0,864628821	0,986899563	0,92139738	0,995633188	0,995633188	0,96069869	1	0,991266376
Students with less than 5	7	31	3	18	1	1	9	0	2
Percentage	0,030567686	0,135371179	0,013100437	0,07860262	0,004366812	0,004366812	0,03930131	0	0,008733624
Not mastered < 5	7	31	3	18	1	1	9	0	2
Percentage Not mastered	0,030567686	0,135371179	0,013100437	0,07860262	0,004366812	0,004366812	0,03930131	0	0,008733624
Average 5<6	4	28	10	42	13	2	11	3	13
Percentage Average	0,017467249	0,122270742	0,043668122	0,183406114	0,056768559	0,008733624	0,048034934	0,013100437	0,056768559
Mastered > 6	218	170	216	169	215	226	209	226	214
Percentage Mastered	0,951965066	0,742358079	0,943231441	0,737991266	0,938864629	0,986899563	0,912663755	0,986899563	0,934497817

Source: Authors

However, particular attention is needed for language subjects, especially Arabic and French, where a notable number of students score below 5. For instance, in French, 31 students scored below this threshold. This is reinforced by the highest non-mastery percentage in French (13.54%) among all subjects examined.
Regarding average scores, students generally display high performance, with particularly high scores in Artistic Education (8.22). However, French has the lowest average score among the subjects analyzed, with an average of 6.96, indicating a potential need for improvement in this discipline.
In terms of skill mastery, most students demonstrate the required competencies in all subjects. However, disparities are observed, with higher mastery rates in Scientific Activity (91.27%) and Physical Education (98.69%), compared to subjects like French and Mathematics, where mastery percentages are lower.
Targeted strategies can be developed to improve skill mastery in these areas, ensuring balanced and comprehensive success for all students.

- **Performance of 5th year**

Table 5: performance of 5th year

	Arabic	French	Social sciences	Mathematics	Islamic education	Physical education	Scientific activity	Art education	Average
Observed average	6,25	5,82	6,44	5,66	7,06	7,29	6,60	7,12	6,53
Number of notes > 5	181	160	174	162	204	236	211	232	209
Percentage	0,77	0,68	0,74	0,69	0,86	1	0,89	0,98	0,89
Number of scores < 5	55	76	62	74	32	0	25	4	27
Percentage	0,23	0,32	0,26	0,31	0,14	0	0,11	0,02	0,11
Checking skills	<5								
Percentage not controlled	0,23	0,32	0,26	0,31	0,14	0	0,11	0,02	0,11
Average 5<6	50	55	36	53	28	4	45	10	53
Average percentage	0,21	0,23	0,15	0,22	0,12	0,02	0,19	0,04	0,22
Controlled >6	131	105	138	109	176	232	166	222	156
Controlled percentage	0,56	0,44	0,58	0,46	0,75	0,98	0,70	0,94	0,66

Source: Authors

Table 5offers a comprehensive analysis of the performance of 5th-grade students across various subjects, along with their distribution according to levels of skill mastery.
Upon examining the observed averages, there is a noticeable variation in scores from one subject to another, ranging from 5.66 to 7.29, with an overall average of 6.53. This suggests significant differences in student performance across different subjects.
The number of students who scored above 5 in each subject is provided, with totals ranging from 160 to 236. The corresponding percentages indicate the proportion of students who successfully achieved scores above 5, offering an insight into the relative performance of students in each field of study.
Similarly, the number of students who scored below 5 is indicated, with totals ranging from 0 to 76 across different subjects. The corresponding percentages highlight the proportion of students who failed to achieve scores above 5, underscoring areas where improvements might be necessary.

- **Performance of 4th year**

Table 6: performance of 4 th year

	Arabe	French	Social sciences	Mathematics	Islamic education	Physical education	Scientific activity	Art education	Average
Observed average	5,72	5,76	5,60	5,43	6,16	7,16	5,82	6,62	6,03
Number of notes > 5	136	177	143	157	161	219	177	199	175
Percentage	0,62	0,81	0,65	0,72	0,74	1	0,81	0,91	0,80
Number of scores < 5	83	42	76	62	58	0	42	20	44
Percentage	0,38	0,19	0,35	0,28	0,26	0	0,19	0,09	0,20
Average 5<6	44	108	50	66	36	1	55	30	74
Average percentage	0,20	0,49	0,23	0,30	0,16	0,005	0,25	0,14	0,34
Controlled >6	92	69	93	91	125	218	122	169	101
Controlled percentage	0,42	0,32	0,42	0,42	0,57	1	0,56	0,77	0,46

Source: Authors

This table provides an in-depth analysis of the performance of 4th-grade students across various subjects, along with their distribution according to levels of skill mastery.
Upon examining the observed averages in each subject, there is a variability in scores, ranging from 5.43 to 7.16, with an overall average of 6.03. This diversity indicates that students have different levels of success depending on the subject, which may reflect the complexity and diversity of the topics covered.
The number of students who scored above 5 in each subject is provided, with totals ranging from 136 to 219. The corresponding percentages highlight that the majority of students succeeded in scoring above 5, with pass rates ranging from 62% to 100%. This suggests generally satisfactory performance by students in most subjects.
Similarly, the number of students who scored below 5 is indicated, with totals ranging from 0 to 83 across different subjects. The corresponding percentages reveal that the proportion of students who failed to score above 5 is relatively low, ranging from 0% to 38%. However, these results also highlight the areas where some students face difficulties.

- **Performance of 3th year**

The provided data offers a comprehensive analysis of student performance across various subjects, along with their distribution according to levels of skill mastery.
Upon examining the observed averages in each subject, we notice a variability in scores, ranging from 5.42 to 6.89, with an overall average of 6.32. These results suggest generally solid performance by students across all subjects, although some subjects may be more challenging than others.
The number of students who scored above 5 in each subject is provided, with percentages showing that the majority of students achieved satisfactory scores, ranging from 60% to 99%.

ISSN: 2658-8455
Volume 6, Issue 12 (2025), pp. 575-594.

Similarly, the number of students who scored below 5 is indicated, with percentages revealing a relatively low proportion of students failing, ranging from 0% to 40%.

Table 7: performance of 3th year

	Arabic	French	Mathematics	Islamic education	Physical education	Scientific activity	Art education	Average
Observed average	6,09	5,83	5,42	6,89	6,79	6,55	6,69	6,32
Number of notes > 5	163	141	129	207	214	183	214	193
Percentage	0,76	0,66	0,60	0,96	0,99	0,85	0,99	0,90
Number of scores < 5	51	73	85	7	0	31	0	21
Percentage not controlled	0,24	0,34	0,40	0,03	0	0,14	0	0,10
Average 5<6	44	46	37	33	0	34	20	69
Average percentage	0,20	0,21	0,17	0,15	0	0,16	0,09	0,32
Controlled >6	119	95	92	174	214	149	194	124
Controlled percentage	0,55	0,44	0,43	0,81	0,99	0,69	0,90	0,58

Source: Authors

Therefore, it would be wise to analyze in more detail the reasons behind these performance differences and implement targeted support measures to help students where they need it most.

- **Performance of 2th year**

Table 8: performance of 2th year

	Arabic	French	Mathematics	Islamic education	Physical education	Scientific activity	Art education	Average
Observed average	6,08	6,19	6,39	6,54	6,95	6,61	6,43	6,45
Number of notes > 5	147	156	161	166	191	170	189	169
Percentage	0,75	0,80	0,83	0,85	0,98	0,87	0,97	0,87
Number of scores < 5	48	39	34	29	4	25	6	26
Percentage not controlled	0,25	0,20	0,17	0,15	0,02	0,13	0,03	0,13
Average 5<6	41	40	24	23	20	25	43	38
Average percentage	0,21	0,21	0,12	0,12	0,10	0,13	0,22	0,19
Controlled >6	106	116	137	143	171	145	146	131
Controlled percentage	0,54	0,59	0,70	0,73	0,88	0,74	0,75	0,67

Source: Authors

The provided data offers a detailed insight into students' performance across different subjects and their distribution according to levels of skill mastery.
The observed average varies slightly from one subject to another, ranging from 6.08 to 6.95, with an overall average of 6.45. These variations likely reflect the complexity of the topics covered and the diversity of skills required to succeed in each subject.
The majority of students scored above 5 in each subject, with percentages ranging from 75% to 98%. However, despite these high scores, a notable number of students scored below 5 in certain subjects, ranging from 2% to 25%. This underscores the importance of closely monitoring student performance and providing additional support where necessary.
Upon examining the distribution of skill mastery, we observe that the percentage of students classified in the non-mastered category (<5) is significantly lower than those in the average (5<6) and mastered (>6) categories. This suggests a general trend towards above-average performance, which is encouraging.
Students with scores between 5 and 6 represent a moderate proportion, with percentages ranging from 10% to 22%. Similarly, students who scored above 6 are also numerous, representing between 54% and 88% of the student population. These figures indicate strong competence in

many areas but also highlight the need to continue encouraging and supporting students to reach their full potential.

- **Performance of 1th year**

Table 9: performance of 1th year

	Arabic	French	Mathematics	Islamic education	Physical education	Scientific activity	Art education	Average
Observed average	5,67	6,18	7,22	6,04	6,29	6,62	6,52	6,36
Number of notes > 5	131	164	167	147	193	167	181	162
Percentage	0,66	0,83	0,84	0,74	0,97	0,84	0,91	0,82
Number of scores < 5	67	34	31	51	5	31	17	36
Percentage	0,34	0,17	0,16	0,26	0,03	0,16	0,09	0,18
Checking skills	Non contrôlé <5							
Percentage not controlled	0,34	0,17	0,16	0,26	0,03	0,16	0,09	0,18
Average 5<6	36	34	21	34	30	20	23	33
Average percentage	0,18	0,17	0,11	0,17	0,15	0,10	0,12	0,17
Controlled >6	95	130	146	113	163	147	158	129
Controlled percentage	0,48	0,66	0,74	0,57	0,82	0,74	0,80	0,65

Source: Authors

Table 9 provides an overview of students' performance across different subjects, as well as their distribution according to levels of skill mastery.

The observed average, ranging from 5.67 to 7.22 with a general average of 6.36, indicates variability in performance across subjects, which is expected given the diversity of topics covered.

It is encouraging to note that in all subjects, a high percentage of students scored above 5, ranging from 66% to 97%. This reflects a relatively high overall level of success among students.

However, the percentage of students scoring below 5, although lower, ranges from 3% to 34%, highlighting the need to understand and address the difficulties faced by some students.

Approximately 18% to 26% of students were not deemed proficient in certain subjects, indicating areas where improvements are needed to ensure all students meet required standards.

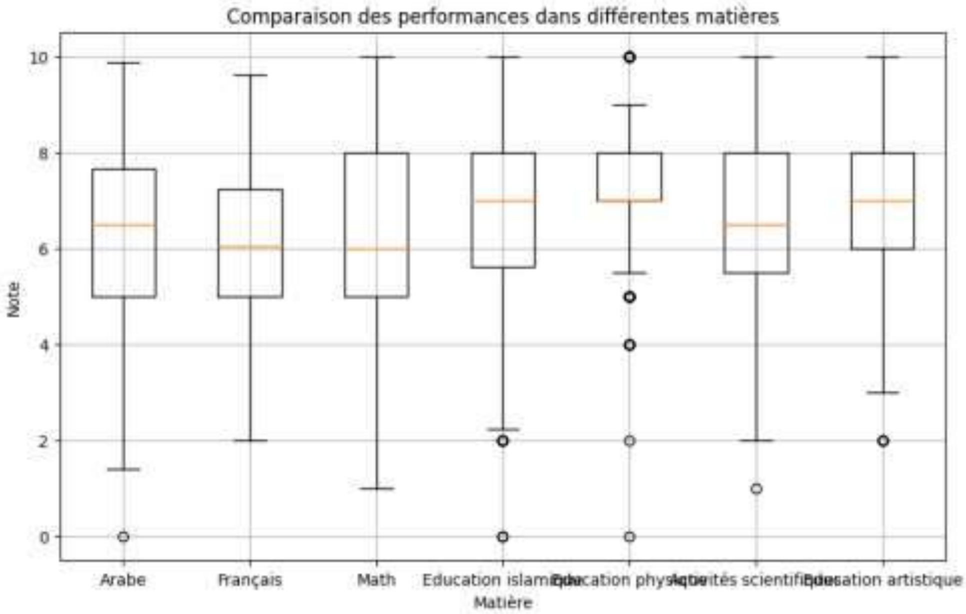
The distribution of students according to levels of skill mastery reveals that 10% to 17% have an average score (between 5 and 6), while 48% to 82% scored above 6, demonstrating a high level of competence among many students.

4.3. A Comparative Performance Analysis

To address the question of whether there are significant differences in mastery of core skills across different subjects taught at ELMOSALA School, we conducted comparative analyses. The box plot summarizes the findings.

When examining the Arabic subject, the median score is around 6. The relatively wide dispersion of scores indicates significant variability in student performance. The low outliers suggest that some students are facing significant difficulties in this subject. For French, the median is slightly lower than that of Arabic, also accompanied by a wide distribution of scores and low outliers. This indicates potential challenges for some students in this subject. Mathematics displays a slightly higher median, around 7, with similarly wide dispersion. The presence of low outliers suggests significant variability, but potentially less pronounced than in other subjects. Islamic Education shows a median around 7.5 with a tighter box, indicating less variability in scores. However, the presence of several low outliers suggests difficulties for some students. Subjects such as Physical Education and Scientific Activities exhibit trends similar to those observed in other subjects. The median score is higher for Scientific Activities, suggesting better overall performance in this subject. Comparing subjects reveals significant differences. For example, subjects like Mathematics, Islamic Education, and Scientific Activities have higher medians compared to other subjects like French and Physical Education.

Fig 2 : Comparison of the performance



Source: Authors

This indicates that students tend to perform better in these subjects. To verify this observation, we conducted the ANOVA test. The obtained ANOVA test statistic is 77.058, indicating the proportion of total variance explained by differences between the compared groups compared to variance due to error or randomness. A high value of this statistic suggests large differences between group means. The p-value associated with this test is 3.339703957965963e-94, an extremely small value, well below common significance thresholds such as 0.05 or 0.01. Such a low p-value indicates that there is an almost zero probability that the observed differences between group means are due to chance. Since the p-value is extremely low, we reject the null hypothesis, which states that all group means are equal. This means that there are significant differences between the means of student scores in different subjects. In other words, student performances vary significantly from one subject to another.

4.4. Correlation analyze

The correlations reveal several important relationships among the variables describing students' characteristics and academic performance. Family-related variables such as the number of siblings, hours of extra classes, and hours of home revision show very weak associations with academic outcomes, as all coefficients remain close to zero. This suggests that these demographic factors do not substantially differentiate academic performance within this sample. By contrast, the strongest correlations are observed among the academic subjects themselves. Arabic, French, and Mathematics are strongly interrelated, with coefficients of 0.75, 0.80, and 0.64 respectively, indicating that students who perform well in one fundamental subject tend to perform well in the others. Islamic Education also exhibits strong correlations with core subjects (e.g., r = 0.82 with Arabic; r = 0.62 with French), suggesting that general cognitive or academic engagement patterns may underlie consistent performance across disciplines. Physical Education shows weaker but still meaningful associations with other subjects (r ranging from 0.38 to 0.70), reflecting its less cognitively demanding nature. Scientific Activities and Artistic Education similarly correlate moderately to strongly with core subjects, reinforcing the idea that overall academic ability plays a central role in shaping student outcomes.

Table 10 correlation analyze

Variables	N. frères	H. suppl.	H. révision	Arabe	Français	Math	Educ. isl.	Educ. phys.	Activ. sci.	Educ. art.	Moyenne	En difficulté
N. frères	1.00	0.07	0.03	0.00	0.03	0.02	-0.02	-0.02	0.00	0.01	0.01	0.00
Heures supplémentaires	0.07	1.00	0.08	-0.03	-0.05	-0.03	-0.02	0.00	-0.05	0.03	-0.03	0.03
Heures de révision	0.03	0.08	1.00	0.02	0.02	0.02	0.02	0.04	0.02	0.03	0.02	-0.04
Arabe	0.00	-0.03	0.02	1.00	0.75	0.64	0.82	0.59	0.72	0.70	0.92	-0.59
Français	0.03	-0.05	0.02	0.75	1.00	0.80	0.62	0.45	0.76	0.50	0.86	-0.50
Math	0.02	-0.03	0.02	0.64	0.80	1.00	0.49	0.38	0.74	0.47	0.81	-0.47
Éducation islamique	-0.02	-0.02	0.02	0.82	0.62	0.49	1.00	0.48	0.59	0.59	0.82	-0.55
Éducation physique	-0.02	0.00	0.04	0.59	0.45	0.38	0.48	1.00	0.44	0.70	0.66	-0.36
Activités scientifiques	0.00	-0.05	0.02	0.72	0.76	0.74	0.59	0.44	1.00	0.53	0.85	-0.55
Éducation artistique	0.01	0.03	0.03	0.70	0.50	0.47	0.59	0.70	0.53	1.00	0.75	-0.44
Moyenne générale	0.01	-0.03	0.02	0.92	0.86	0.81	0.82	0.66	0.85	0.75	1.00	-0.61
En difficulté (0/1)	0.00	0.03	-0.04	-0.59	-0.50	-0.47	-0.55	-0.36	-0.55	-0.44	-0.61	1.00

Source: Authors

The average grade (Moyenne générale) is highly correlated with all subject-specific scores, particularly Arabic ($r = 0.92$), French ($r = 0.86$), and Mathematics ($r = 0.81$), confirming that these disciplines contribute heavily to the overall evaluation system. Conversely, the variable "En difficulté" (0 = no difficulty, 1 = difficulty) is strongly negatively correlated with performance indicators, especially with Arabic ($r = –0.59$), Scientific Activities ($r = –0.55$), and the overall average ($r = –0.61$). This pattern suggests that academic difficulty is closely linked to broad underperformance across subjects rather than to deficits in isolated areas. The consistently strong negative correlations imply that the classification of students as "in difficulty" is well aligned with their academic results and accurately reflects their overall performance profile.

Overall, the correlation matrix highlights that academic subjects cluster strongly together, forming a coherent construct of general academic achievement, while socio-familial variables exert minimal influence. These insights justify the use of multivariate models focusing primarily on academic predictors and underscore the need for intervention strategies that target global learning challenges rather than isolated subject-specific issues.

4.5. Analysis of Influential Factors:

To address the question of which demographic, familial, environmental, and institutional factors influence the academic performance of students at ELMOSALA School, we conducted a multiple regression analysis. The detailed results of this analysis are presented in the table 11 below.

The results of this multiple regression analysis reveal several important insights into the factors influencing the academic performance of students at ELMOSALA School, based on the variables included in the model.

Firstly, let's examine the variables that did not demonstrate statistical significance (p-value > 0.05). Variables x1 to x9, which include aspects such as gender, parents' educational level,

ISSN: 2658-8455
Volume 6, Issue 12 (2025), pp. 575-594.

parents' marital status, housing arrangement, number of siblings, academic support, extra hours, and number of hours of revision at home, appear to have no significant impact on students' academic performance in this model.

Table 11: linear regression results

```
                            OLS Regression Results
==============================================================================
Dep. Variable:                      y   R-squared:                       0.995
Model:                            OLS   Adj. R-squared:                  0.995
Method:                 Least Squares   F-statistic:                 1.177e+04
Date:                Wed, 01 May 2024   Prob (F-statistic):               0.00
Time:                        20:05:16   Log-Likelihood:                 930.29
No. Observations:                1012   AIC:                            -1827.
Df Residuals:                     995   BIC:                            -1743.
Df Model:                          16
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
x1             0.0111      0.006      1.754      0.080      -0.001       0.024
x2             0.0019      0.004      0.504      0.614      -0.006       0.009
x3            -0.0003      0.003     -0.084      0.933      -0.006       0.006
x4             0.0069      0.009      0.745      0.456      -0.011       0.025
x5             0.0046      0.007      0.637      0.525      -0.010       0.019
x6             0.0004      0.004      0.118      0.906      -0.007       0.007
x7             0.0069      0.007      0.973      0.331      -0.007       0.021
x8             0.0177      0.007      2.710      0.007       0.005       0.030
x9            -0.0062      0.004     -1.758      0.079      -0.013       0.001
const         -0.0531      0.024     -2.247      0.025      -0.099      -0.007
x10            0.1731      0.004     40.064      0.000       0.165       0.182
x11            0.1245      0.004     30.011      0.000       0.116       0.133
x12            0.1449      0.003     53.612      0.000       0.140       0.150
x13            0.1533      0.003     51.791      0.000       0.147       0.159
x14            0.1369      0.004     35.673      0.000       0.129       0.144
x15            0.1419      0.003     47.500      0.000       0.136       0.148
x16            0.1307      0.004     32.726      0.000       0.123       0.138
==============================================================================
Omnibus:                      264.274   Durbin-Watson:                   2.115
Prob(Omnibus):                  0.000   Jarque-Bera (JB):             8718.183
Skew:                          -0.489   Prob(JB):                         0.00
Kurtosis:                      17.346   Cond. No.                         142.
==============================================================================

Notes:
[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.
```

Source: Authors

On the other hand, variables x10 to x16, representing the different subjects studied such as Arabic, French, Mathematics, Islamic Education, Physical Education, Scientific Activities, and Art Education, all show a significant and positive impact on academic performance. The high coefficients and p-values close to zero indicate that performance in these subjects is strongly correlated with students' overall average.

4.6. Identification of Predictive Characteristics:

To answer the question of which characteristics are the most predictive of academic success among demographic, familial, environmental, and institutional variables, advanced data analysis techniques such as Principal Component Analysis (PCA) were employed to identify the most predictive features of academic success among the studied variables. The table summarizes the PCA results.

Table 12: Principal components analysis (PCA

Variable	PC1	PC2	PC3	PC4	PC5
Gender	0.993855	0.088345	0.047055	-0.006833	0.007149
Father's level of education	0.000932	0.022188	0.037075	-0.001234	-0.021025
Mother's level of education	-0.000677	-0.007596	0.044181	0.345458	0.892641
Parental situation (separated or not)	-0.000930	-0.003940	0.052753	0.912351	-0.325374
Housing layout	-0.000190	0.000817	0.001657	0.014341	0.045554
Number of brothers	0.000850	-0.000135	-0.000780	-0.027121	-0.020546
Tutoring	0.002152	0.002946	0.044366	-0.056144	0.036081
Overtime	0.000081	-0.002067	-0.000335	-0.091436	-0.086523
Number of hours of revision at home	0.000714	0.004556	-0.005828	0.060152	-0.108941
Phone use	0.001387	-0.008054	-0.023057	0.004839	-0.189827
Arabic	-0.000000	6.018531e-36	-0.000000	0.000000	-2.168404e-19
French	0.052657	-0.386184	-0.287269	0.003593	-0.008216
Math	0.018521	-0.347218	0.190439	-0.049169	-0.003581
Physical education	0.011999	-0.471107	0.611653	-0.078105	-0.063860
Islamic education	0.046856	-0.375724	-0.635388	-0.034695	0.060506
Scientific activities	0.057153	-0.141418	-0.089667	0.094632	-0.110757
Art education	0.019415	-0.418139	0.185214	0.044560	0.106552
Average	0.042854	-0.214654	-0.217833	0.106360	-0.077705
Difficulties(yes/no)	0.036144	-0.343229	-0.052111	0.011931	-0.012875

Source: Authors

The loadings of the principal components (PCs) indicate the relative importance of each original variable in each principal component.

For the first principal component (PC1), the variable 'Sex' with a loading of 0.993855 is by far the most dominant. This variable almost entirely explains this component, while the other variables have very low coefficients. This suggests that 'Sex' is one of the most predictive variables of the data variance.

In the second principal component (PC2), the variables 'Physical Education' (-0.471107), 'Art Education' (-0.418139), and 'Math' (-0.347218) are the most influential, with significant negative coefficients. These characteristics play an important role and capture specific variations in the data not explained by PC1.

The third principal component (PC3) shows that 'Physical Education' (0.611653) and 'Islamic Education' (-0.635388) have high but opposite coefficients, indicating an inverse relationship. PC3 captures variations where these two variables are inversely correlated, revealing a different dimension of variance in the data.

For the fourth principal component (PC4), the variables 'Parental Status (Separated or Not)' (0.912351) and 'Mother's Educational Level' (0.345458) have high coefficients, especially parental status. PC4 is dominated by this variable, with a significant contribution from the mother's educational level, suggesting a dimension mainly related to these variables.

The fifth principal component (PC5) shows high and inverse coefficients for 'Mother's Educational Level' (0.892641) and 'Parental Status (Separated or Not)' (-0.325374). This component captures a dimension where these two variables have an inverse relationship, indicating another important variation in the data.

Overall, 'Sex' is by far the most influential in the first principal component, suggesting that it is the most predictive variable of the total variance in the data. The variables 'Math', 'Physical Education', and 'Art Education' are important in the subsequent principal components, capturing additional dimensions of variance. The 'Mother's Educational Level' and 'Parental Status (Separated or Not)' are significant in the principal components, indicating complex relationships between them.

ISSN: 2658-8455
Volume 6, Issue 12 (2025), pp. 575-594.

For a more in-depth analysis, it would be useful to visualize the loadings of the principal components to better understand the relationships between the original variables and the PCs, to examine the scores of the principal components to identify the most influential observations in each PC, and to apply modeling techniques (regression, classification) using the PCs to assess their ability to predict academic success. This could help confirm the importance of the identified variables and refine predictions.

4.7. Model Predictive Evaluation:

To address the question of which machine learning models are most effective for predicting academic performance of students at ELMOSALA school, we constructed and evaluated several predictive models using different machine learning algorithms such as decision trees, neural networks, support vector machines (SVM), and ensemble methods. The table below summarizes the results of cross-validation:

Table 13: Evaluation of predictive models

	Accuracy	F1 Score	RMSE	R-squared
Logistic regression	98.02%	96.17%	-	-
Decision tree	92.78%		2.062	-1.034
KNN	96.05%	91.14%	2.027	-0.965
Random Forest	-	-	2.048	-1.006
Support Vector Machine	-	-	0.188	-1.013

Source: Authors

Interpreting the results of different machine learning models for predicting academic performance of students at ELMOSALA school reveals several key findings.
Logistic Regression exhibits the highest precision among all evaluated models, with an Accuracy of 98.02% and an F1 Score of 96.17%. These results indicate that this model is capable of accurately predicting academic performance of students at the school.
On the other hand, Decision Tree, although showing a respectable precision of 92.78%, presents less convincing performances in terms of RMSE (2.062) and R-squared (-1.034), suggesting difficulties in generalizing the data.
The KNN model also demonstrates good results with a precision of 96.05% and an F1 Score of 91.14%. However, its RMSE (2.027) and R-squared (-0.965) values indicate a slight tendency towards overfitting.
Lastly, the SVM model stands out for its low RMSE (0.188) and R-squared (-1.013), suggesting an ability to provide precise predictions with minimal error.
In summary, logistic regression appears to be the safest choice for predicting academic performance of students at ELMOSALA school, followed by SVM and KNN. However, it is crucial to consider all aspects of model performance to obtain a comprehensive assessment of their effectiveness in this specific context.

5. Discussion

The results of the empirical analysis confirm several theoretical findings highlighted in the literature. First, the strong and significant correlations between subject-specific grades and overall academic performance align with the work of Arnold & Pistilli (2012) and Burgos et al. (2018), who emphasized the predictive value of academic engagement and early performance indicators. Similarly, the regression results showing that socio-economic status and parental education significantly influence student achievement reinforce the conclusions of Rizvi et al. (2019) and UNESCO (2020), suggesting that academic success is shaped not only by cognitive factors but also by broader structural inequalities. This convergence between empirical results

and existing research strengthens the validity of our findings and demonstrates the relevance of integrating contextual variables into educational analysis.
Moreover, the positive effect of active learning variables (e.g., subject mastery) echoes the findings of Romero & Ventura (2013) and Huang & Fang (2013), who underscored the role of social interaction and engagement in learning processes. Although our data do not explicitly measure interaction in virtual environments, the significance of academic indicators suggests that consistent participation and mastery remain central components of academic success in both physical and digital learning settings.
At a methodological level, the comparison of machine learning models with traditional statistical methods offers insights consistent with previous studies such as Okubo et al. (2017) and Jha et al. (2019). While our predictive models confirm that algorithms like Random Forest and SVM can outperform linear methods in accuracy, the regression analysis remains essential for interpretability—a limitation repeatedly emphasized in the literature. This dual approach thus reflects the tension highlighted by Karimi et al. (2020) between predictive power and model transparency. Our findings contribute to this debate by showing that a combination of interpretable models and advanced predictive techniques offers the most informative framework for understanding academic performance.
Overall, the empirical results not only support existing theoretical frameworks but also highlight the importance of integrating institutional context into the analysis. Differences in subject performance, the role of parental background, and the predictive potential of machine learning models all suggest that academic achievement is a multidimensional construct requiring both quantitative sophistication and contextualized interpretation. The alignment between theoretical claims and empirical evidence reinforces the robustness of the study while providing practical implications for school administrators seeking to identify at-risk students and design targeted interventions.

6. Conclusion

After examining in detail the academic performances of students at ELMOSALA school, along with existing research on improving these performances, it is clear that significant progress can be made by better understanding the specificities of each discipline and adapting pedagogical strategies accordingly.
The findings of this study underscore the importance of considering a variety of factors, ranging from demographic and familial characteristics to the subjects studied and the quality of teachers, in designing programs to enhance academic performance. Moreover, the use of analytical approaches such as principal component analysis and logistic regression can provide valuable insights into understanding the key factors influencing students' academic success.
To advance in this field, it is essential that the developed models are understandable and usable by education practitioners, to enable effective and targeted intervention. Additionally, a more thorough integration of contextual aspects specific to each institution or educational system would strengthen the relevance and applicability of research in different educational environments.
In conclusion, this study provides a solid foundation for the development of targeted educational strategies aimed at improving the academic performance of students at ELMOSALA school. By combining insights from this comprehensive analysis with the findings of existing research, it is possible to create effective and tailored programs, thereby fostering student success in a diverse school environment.

Références

(1). Arnold, K. E., & Pistilli, M. D. (2012). Course signals at Purdue : Using learning analytics to increase student success. *Proceedings of the 2nd International Conference on Learning Analytics and Knowledge*, 267-270. https://doi.org/10.1145/2330601.2330666

(2). Baker, R. S. J. d, & Yacef, K. (2009). The State of Educational Data Mining in 2009 : A Review and Future Visions. *Journal of Educational Data Mining*, *1*(1), Article 1. https://doi.org/10.5281/zenodo.3554657

(3). Barber, R., & Sharkey, M. (2012). Course correction : Using analytics to predict course success. *ACM International Conference Proceeding Series*. https://doi.org/10.1145/2330601.2330664

(4). Burgos, C., Campanario, M. L., Peña, D. de la, Lara, J. A., Lizcano, D., & Martínez, M. A. (2018). Data mining for modeling students' performance : A tutoring action plan to prevent academic dropout. *Computers & Electrical Engineering*, *66*, 541-556. https://doi.org/10.1016/j.compeleceng.2017.03.005

(5). Calvo-Flores, M., Gibaja, E., Pegalajar Jiménez, M. del C., & Pérez, O. (2006). Predicting Students' Marks from Moodle Logs using Neural Network Models. *Current Developments in Technology-Assisted Education*.

(6). Carine, T. L. (2016). *Améliorer la qualité de l'éducation de base pour tous : Des recommandations internationales aux politiques nationales*.

(7). Haji, R. (2011). *Éducation, croissance économique et développement humain : Le cas du Maroc*.

(8). Hassan, M., Bin Alam, Md. S., & Ahsan, T. (2018). Emotion Detection from Text Using Skip-thought Vectors. *2018 International Conference on Innovations in Science, Engineering and Technology (ICISET)*, 501-506. https://doi.org/10.1109/ICISET.2018.8745615

(9). Huang, Y.-M., Liang, T.-H., Su, Y.-N., & Chen, N.-S. (2012). Empowering personalized learning with an interactive e-book learning system for elementary school students. *Educational Technology Research and Development*, *60*(4), 703-722. https://doi.org/10.1007/s11423-012-9237-6

(10). Jha, P., Jha, R., & Sharma, A. (2019). *Behavior Analysis and Crime Prediction using Big Data and Machine Learning*. *8*(1).

(11). Karimi, H., Derr, T., Huang, J., & Tang, J. (2020). *Online Academic Course Performance Prediction using Relational Graph Convolutional Neural Network*.

(12). Kotsiantis, S., Tselios, N., Filippidi, A., & Komis, V. (2013). Using learning analytics to identify successful learners in a blended learning course. *International Journal of Technology Enhanced Learning*, *5*(2), Article 2. https://doi.org/10.1504/IJTEL.2013.059088

(13). Okubo, F., Yamashita, T., Shimada, A., & Ogata, H. (2017). A neural network approach for students' performance prediction. *Proceedings of the Seventh International Learning Analytics & Knowledge Conference*, 598-599. https://doi.org/10.1145/3027385.3029479

(14). Pardos, Z. A., Heffernan, N. T., Anderson, B., & Heffernan, C. L. (2007). The Effect of Model Granularity on Student Performance Prediction Using Bayesian Networks. Dans C. Conati, K. McCoy, & G. Paliouras (Éds.), *User Modeling 2007* (p. 435-439). Springer. https://doi.org/10.1007/978-3-540-73078-1_60

(15). Rizvi, S., Rienties, B., & Khoja, S. (2019). The role of demographics in online learning; A decision tree based approach. *Computers & Education*, *137*. https://doi.org/10.1016/j.compedu.2019.04.001

(16). Roberge, D., Rojas, A., & Baker, R. (2012). Does the length of time off-task matter? *ACM International Conference Proceeding Series*. https://doi.org/10.1145/2330601.2330657

(17). Romero, C., & Ventura, S. (2013). Data mining in education. *WIREs Data Mining and Knowledge Discovery*, *3*(1), Article 1. https://doi.org/10.1002/widm.1075

(18). Sobhi, T., Sophie, C., & Amapola, A. (2010). *Education au Maroc : Analyse du secteur—UNESCO Bibliothèque Numérique*. https://unesdoc.unesco.org/ark:/48223/pf0000189743

(19). Wang, S.-P., & Chen, Y.-L. (2018). Effects of Multimodal Learning Analytics with Concept Maps on College Students' Vocabulary and Reading Performance. *Educational Technology and Society*, *21*, 12-25.