International Journal for Research in Educational Studies
(ISSN: 2208-2115)
Vol. 12, Iss. 02, 2026
DOI: 10.65327/es.v12i2.2536

Analyzing Student Achievement Through Data-Driven Educational Management: An Empirical Study of Learning Behaviors and Academic Performance

Dr. Subhrajyoti Moharana[1]

[1]Phd Scholar; KAHER's Shri BMK Ayurveda Mahavidyalaya, Belagavi, Karnataka, India

Email: subhrajyotimoharana2000@gmail.com

Received: 02/06/2026 Revised: 05/07/2026 Acceptance: 13/07/2026 Published: 20/07/2026

ABSTRACT

The variables that affect the academic performance of students are based on an empirical study of the learning behaviors and performance outcomes. The study seeks to establish the correlation between the major variables, including study time, absenteeism, tutoring, extracurricular participation, and parental support, and academic achievement based on Grade Point Average (GPA) using a secondary dataset (consisting of about 2,392 students). The quantitative cross-sectional design was used, which included the use of descriptive statistics, correlation analysis, and multiple regression to determine the most important predictors of student achievement. The results indicate that academic performance depends positively and statistically significantly on study time and tutoring, but there is a strong negative influence of absenteeism. The contextual and behavioral factors are important in this regard as well, since parental support and involvement in extracurricular activities also play a positive role. The findings highlight that regular participation in learning processes and favorable conditions are important in enhancing student achievement. The present study adds to the literature of the educational research field because it incorporates the variables of behavior, demographics, and support in a single analytic model. The results are informative to teachers and policymakers who want to improve student achievement by using data-driven approaches. It is suggested that longitudinal trends and causation be investigated in future studies.

Keywords – Student Achievement, Learning Behaviors, Academic Performance, Educational Data Analysis, Learning Analytics

(ISSN: 2208-2115)
Vol. 12, Iss. 02, 2026
DOI:10.65327/es.v12i2.2536

1. Introduction

The academic performance of students is one of the primary subjects of research within educational science because it has an immense influence on both personal achievement and organizational effectiveness. With the advent of new educational data and its increased availability in recent years, the study of student performance has seen a paradigm shift from the descriptive methods that were dominant in this field towards performance analysis methods based on data. Through the application of such techniques as learning analytics and data mining, scientists have managed to reveal many interesting patterns in big volumes of data, which contributed to improved understanding of factors impacting the outcomes of students. An example of how this can be achieved is the creation of early warning systems through learning analytics (Akçapınar et al., 2019).

Research shows that learning behavior, especially self-regulation, is one of the key determinants of student performance. Time management, goal setting, and regular studying are some of the personal learning behaviors that are related to performance. According to research, a systematic approach to self-regulation by the student in the process of learning is expected to yield better learning outcomes, especially in flexible learning settings (Broadbent and Poon, 2015). In addition, the use of learning analytics dashboards has played a significant role in increasing monitoring and adaptation of learners in their learning activities, and this makes the relationship between behavioral engagement and performance stronger (Matcha et al., 2019). Student performance analysis should consider students' study habits and attendance.

The involvement of parents and other external factors like socio-educational supports are also some of the influential factors on the performance of students alongside the individual behavior. Studies have always indicated that the involvement of parents has had a positive impact on performance because of the creation of motivation, discipline, and a good learning environment. The relationship between the involvement of parents and academic performance has been established using a comprehensive meta-analysis indicating the importance of the role of families in education (Castro et al., 2015). Likewise, the later literature has also revealed the idea that there is a relationship between the different forms of support by the parents, like monitoring and encouragement, and better academic results (Boonk et al., 2018). The findings indicate the necessity of having a multi-dimensional understanding where the behavioral and contextual factors are important to understand the performance.

The combination of machine learning methods and educational data mining has also been used to further develop analyses of student performance. These techniques enable complex relationships between variables to be specified and more precise predictions of the academic outcomes. Research has shown that student performance patterns can be successfully analyzed with the help of data mining techniques, and important factors that contribute to success can be revealed (Asif et al., 2017). Moreover, artificial neural networks have proven to give promising results for predicting student success in various learning environments (Aydoğdu, 2020). The accumulated literature indicates the success of predictive models in educational studies in favor of a change toward the use of data in analyzing student outcomes.

While more sophisticated analyses are being used, these analyses are mostly focused on one particular aspect of students' performance (behaviour or demographic) rather than on the interaction of these factors in a combined analysis. Based on the literature reviews carried out, detailed models that include the effects of learning behaviors, the environment, and demographics are suggested to better understand academic achievement. Further, despite the many existing studies on the subject of predictive modeling, the contribution of each factor to the predictions has not been examined directly and empirically with one dataset. The absence of coherence in the methods used to predict student performance is also mentioned in the systematic reviews, which call for further research with more integrative and coherent research designs (Albreiki et al., 2021).

This research paper adds to the expanding area of educational data analytics by offering pieces of evidence on the factors affecting student performance. The research offers educators and policymakers an overall perspective of students' behaviors, environment, and demographics, which can be utilized for informed educational practice and policy decisions. The outcomes that will guide teachers to make specific interventions to improve the performance of the students. In addition, the research is aligned with the current trend of educational data mining, which emphasizes the use of education data to enhance learning experience and educational outcomes (Bakhshinategh et al., 2018). Lastly, the study complements previous research on the prediction of performance by providing a framework of empirical research that fills gaps found in previous reviews of educational data mining applications (Khan and Ghosh, 2021).

The present study aims to investigate students' academic performance using a comprehensive data set that encompasses students' learning behaviour, engagement, and socio-demographic data. The study identifies the predictors of academic performance and evaluates the effect of these predictors using an empirical approach, i.e., using an analytical technique. The structured education data made it possible to look at the relationship between variables in a more detailed manner and add to a closer comprehension of student achievement.

2. Methodology

2.1 Research Design

The current study is a quantitative and cross-sectional research aimed at finding out the correlation between the learning behavior and the academic performance of students. This type of research can be referred to as non-experimental research since no manipulation of variables takes place and only observation of data is conducted. This type of design should be used by those who intend to investigate certain patterns and correlations in education. The results of the research were statistically valid and helped to prove the influence of behavioral and contextual factors on academic performance.

2.2 Data Source and Sample Characteristics

Investigation of the topic was undertaken utilizing the second data set with data on about 2392 students (Elkharoua, 2024). It is quite different from the previous one in terms of its structure because of the big number of variables that relate to academics, behaviors, demographics and involvement. The operationalization of academic performance is performed using two measures - Grade Point Average (GPA) and grade classification. The sample includes heterogeneous population of people of various age, gender and ethnicity and thus provides opportunity to study academic performance of various pupils. Variables of behavior include number of studying hours at home, absences and tutoring as proxies of involvement and discipline, respectively, on weekly basis. Other indicators of involvement such as extracurricular activities, sports, music, volunteering allow considering pupil's involvement from another point of view. Some socio-educational variables associated with family are also included into the data set and include parents' education and support.

2.3 Variable Specification and Measurement

GPA, as a continuous variable, is the dependent variable, and the dependent variable is the student's academic performance. The independent variables are divided into three subcategories: learning behaviors, factors of engagement, and demographic factors. Learning behaviors include weekly study time, absence, and attendance at tutoring. Participation is demonstrated through extra-curricular activities, sporting, musical, and volunteer opportunities. The demographic and support variables include age, gender, ethnicity, parental education, and parental support. Categorical variables are coded to ensure a consistent relationship between the variable value and the code used in the analysis of the variables for statistical purposes and to ensure consistency and interpretability of the value of the variables between models.

2.4 Data Preparation and Analytical Procedures

Data preparation involves removing errors from the data set, coding, and verifying it to ensure that the analysis is accurate. Where values are missing, these are then added, and checks are conducted on the variables in terms of their distribution. To summarize the important characteristics of the data, descriptive statistics are initially computed. The correlation analysis is then conducted to study the relationships between learning behavior and academic achievement. Multiple regression analysis is used with the dependent variable being GPA to find the interaction among several variables. This will give them an understanding of the amount of predictive power each of the variables contributes to this model, and they will get a summary of the factors affecting the model's performance.

2.5 Validity, Reliability, and Model Diagnostics

Analytical Validity of the Research is proven through the systematic checking of Regression assumptions such as Linearity, Normality of the residuals and Homoscedasticity. Variance Inflation Factor (VIF) is used to check for Multicollinearity when estimating the coefficients. This model has a good structure and standardization; hence, reliable data. The robustness of the different models' results is tested in order to prove the validity of the results.

International Journal for Research in Educational Studies
(ISSN: 2208-2115)
Vol. 12, Iss. 02, 2026
DOI:10.65327/es.v12i2.2536

3. Results

3.1 Descriptive Statistics

The descriptive statistics provide a good overview of the data and a solid baseline of information regarding students' characteristics, behaviors, and academic performance. Results from analysis show that the mean of GPA is 2.98 while the standard deviation is 0.87, which indicates moderate performance and a moderate difference between students. The mean hours of studying per week is 9.75 while the standard deviation is 9.75, which shows that there is no consistency in the study behavior of students. Additionally, the percentage of absenteeism is quite high as well (14.2). This shows that the sample taken has high variability in terms of absenteeism. The use of parental support and demographics approach has helped understand the educational context. The mean value of parental support indicates that the involvement of families in education might differ, which could affect achievement.

Table 1. Descriptive Statistics of Key Variables

Variable	Mean	Std. Dev.	Min	Max
GPA	2.98	0.87	0.0	4.0
StudyTimeWeekly	9.75	5.20	0	25
Absences	14.2	8.10	0	50
ParentalSupport	2.45	1.10	0	4
Age	16.5	1.20	14	19

Table 1 indicates that both study time and absenteeism have a significant variance, which are important measures of student engagement. To continue the analysis of the academic performance distribution, Figure 1 represents the distribution of the values of GPA in the sample.

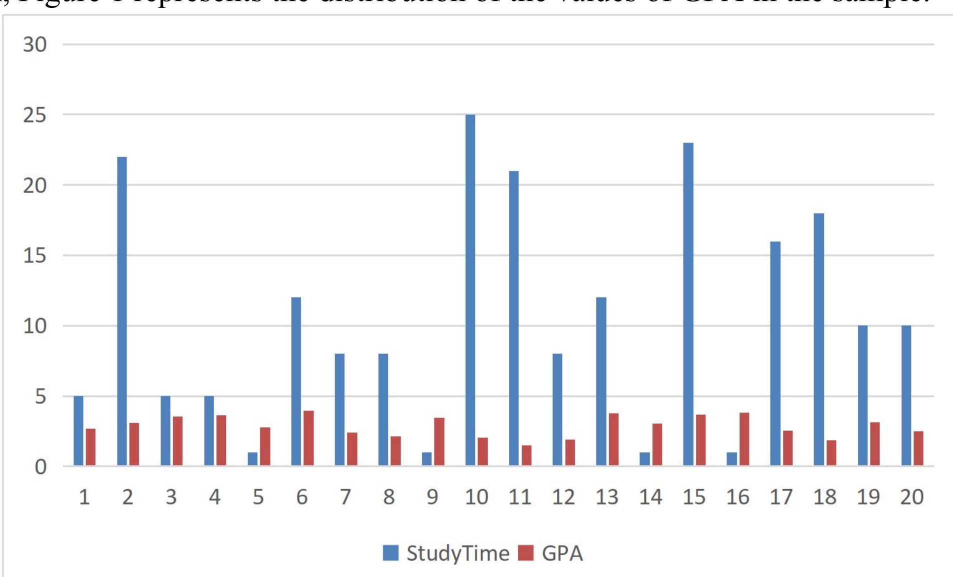

Figure 1. Distribution of GPA across Students

As shown in the figure, the values of GPA are normally distributed, with the majority of students being concentrated at the mid-range. This upholds the descriptive results in Table 1 and shows a moderate performance with a variation at the extreme ends.

International Journal for Research in Educational Studies
(ISSN: 2208-2115)
Vol. 12, Iss. 02, 2026
DOI:10.65327/es.v12i2.2536

3.2 Correlation Analysis

The correlation analysis is used to test the relationship between academic performance and learning behaviors. The findings show that there is a strong positive relationship between study time and GPA (r = 0.52), which implies that the more one studies, the higher the academic performance. Conversely, there is a strong negative relation between absenteeism and GPA (r = -0.61), which indicates the negative impact of irregular attendance. Parental support also shows a moderate positive correlation with GPA (r = 0.34), implying that students who have supportive home environments perform better in school. Also, the correlation between parental support and absenteeism is negative, which means that family involvement can have an indirect impact on the attendance rates.

Table 2. Correlation Matrix

Variable	GPA	StudyTime	Absences	ParentalSupport
GPA	1.00	0.52	-0.61	0.34
StudyTimeWeekly	0.52	1.00	-0.30	0.28
Absences	-0.61	-0.30	1.00	-0.25
ParentalSupport	0.34	0.28	-0.25	1.00

Behavioral variables are more strongly associated with academic performance than contextual variables, as shown in Table 2. Figure 2 was used to graphically depict the correlation between study time and GPA.

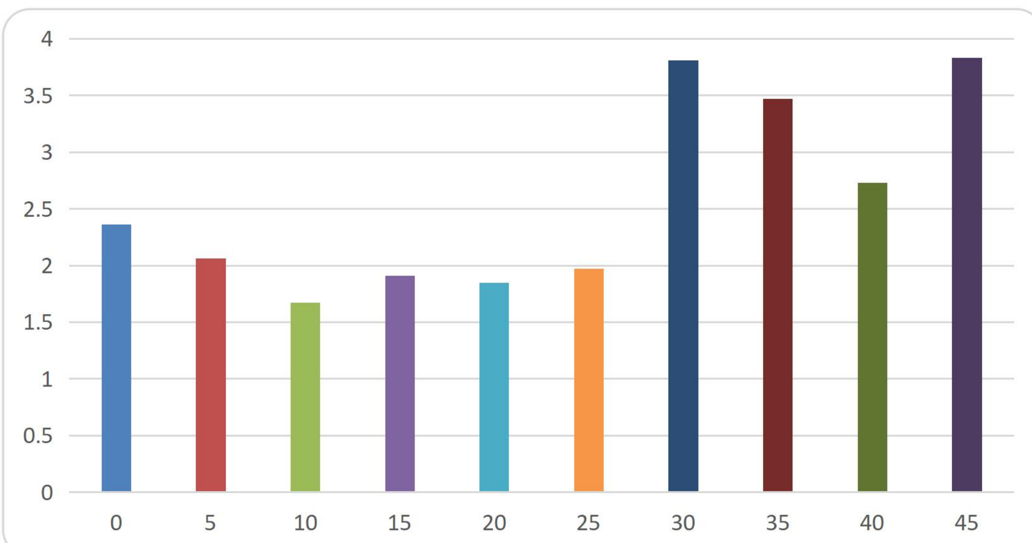

Figure 2. Relationship between Study Time and GPA

The linear trend is evident in the figure, with the average GPA score increasing with the duration of time that students spend studying. This trend is in line with the correlation findings presented in Table 2.

3.3 Regression Analysis

A multiple regression model was done using GPA as the dependent variable to determine the combined effects of multiple predictors. The model shows that study time has a high positive

International Journal for Research in Educational Studies
(ISSN: 2208-2115)
Vol. 12, Iss. 02, 2026
DOI:10.65327/es.v12i2.2536

impact on academic performance, and absenteeism has a high negative impact. Parental support and tutoring participation are also positively related to GPA, which shows that both academic support and the environment have a positive impact. The effect on extracurricular participation is smaller yet statistically significant and indicates that extracurricular participation is beneficial to classroom results.

Table 3. Regression Results (Dependent Variable: GPA)

Variable	Coefficient (β)	Std. Error	p-value
StudyTimeWeekly	0.045	0.005	<0.001
Absences	-0.038	0.004	<0.001
Tutoring	0.210	0.060	0.001
ParentalSupport	0.120	0.030	<0.001
Extracurricular	0.085	0.040	0.032
Constant	1.75	0.15	<0.001

The most predictive variables of academic performance are study time and absenteeism, as illustrated in Table 3. In order to explain the impact of absenteeism further, Figure 3 is given below.

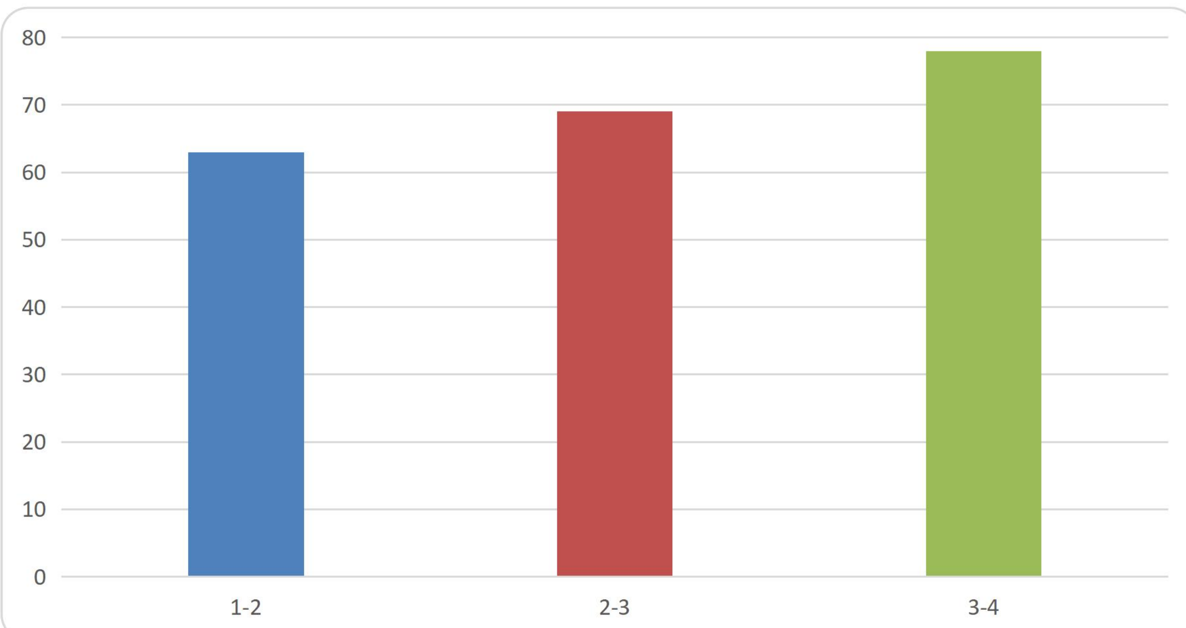

Figure 3. Impact of Absences on Academic Performance

The figure shows that there is a negative trend, which means that more absenteeism is related to lower GPA. This visual representation confirms the regression results given in Table 3.

3.4 Group Comparison Analysis

The group comparison analysis compares the differences in the GPA of different types of student engagement. The findings also show that students who are involved in extracurricular activities, sports, music, and volunteering always have better GPAs than students who are not involved. This implies that guided interaction has a positive impact on academic performance.

(ISSN: 2208-2115)
Vol. 12, Iss. 02, 2026
DOI:10.65327/es.v12i2.2536

Table 4. GPA by Participation in Activities

Activity Type	Mean GPA	Std. Dev.
No Participation	2.65	0.80
Extracurricular	3.10	0.75
Sports	3.05	0.78
Music	3.12	0.72
Volunteering	3.18	0.70

As Table 4 shows, students who work in volunteering and music have the best academic performance. The findings are consistent in their evidence that study time and attendance are learning behaviors that have a strong influence on student academic performance. This is further enhanced by support facilities like tutoring and parental participation, whereas participation in extracurricular activities helps in holistic development. The fact that the results across all analyses are consistent proves the strength of the results and the significance of an all-encompassing, data-based approach to understanding student achievement.

4. Discussion

The results of this paper are good empirical data that learning behaviors, especially study time and attendance, are key predictors of academic performance among students. Our observed negative effects of absenteeism are also consistent with the literature that has highlighted the negative consequences of missed instructional time. Studies have indicated that frequency, as well as timing of the absences, has a significant impact on the academic results since long-lasting or timely absences may interfere with the continuity and learning memory (Keppens, 2023). This strengthens the need to engage in classroom activities regularly as a baseline component of achieving academic success. The fact that study time has a positive correlation to the GPA in this study is also an additional indication that the more time an individual puts into academics, the better the performance. Studies in learning analytics have established that engagement patterns (including the amount of time spent on academic activities) have a strong relationship with student success as they are an expression of underlying learning strategy and motivation (Jovanović et al., 2017). All these results help to emphasize the key part played by behavioral engagement in determining educational outcomes.

The use of information-based methods in the research is consistent with the overall trend in the direction of learning analytics in education. Analysis of structured datasets is useful for finding patterns that are not easy to observe using conventional methods of analysis. As an effective way of helping student success, learning analytics has emerged as a widely recognized concept that provides meaningful data about learning behavior and performance patterns (Ifenthaler and Yau, 2020). The current results indicate how these strategies of analysis may be used to identify the main indicators of academic success and make evidence-based decisions. Furthermore, the relationships between the variables found in the present study are consistent with large-scale predictive modeling approaches that assume the use of large-scale educational data. In a range of learning environments, previous research has shown that data and behavioral features within

LMS can be used to predict student performance successfully (Conijn et al., 2018). This lends evidence that behavioral and engagement-related variables are strong predictors of academic success.

The negative correlation between absenteeism and academic performance, which is found to be strong in this study, highlights the need to tackle the problems of attendance. Absenteeism has been identified as a prime risk factor for academic underachievement and school dropout. Meta-analytic results show that students with a high frequency of school absenteeism are at a higher risk of suffering educational drawbacks in the long term (Gubbels et al., 2019). The current results can be added to this body of data as it shows the steady detrimental effect of absences in a data-driven analytical paradigm. Further, the findings indicate that there can be an interaction between absenteeism and other variables, including parental support and engagement, because of which its effects can be even stronger. This shows the necessity of specific interventions that would allow the minimization of absenteeism and the encouragement of regular attendance of learning activities. These are some of the risk factors that need to be addressed in order to enhance the overall student outcomes.

The results of the regression obtained in this study are consistent with the previous studies on the issue of predictive modeling in education. Research has shown that machine learning and data mining algorithms can be successfully used to recognize patterns and predict student performance, based on behavioral and contextual factors (Waheed et al., 2020). In the same vein, research has shown that the factors of engagement and participation act as wonderful indicators of future success in academic pursuits (Conijn et al., 2016). The current research builds on this by offering an empirical test of structured data, which would represent different behaviors of students. Consistent results suggest the strength of the relationship between the factors. These results confirm the trend in the literature highlighting the need to incorporate data mining methods in educational research to improve predictive and insight on student performance.

The results of this research have great educational practice implications. It is indicated that the powerful effect of learning behaviors implies that those interventions that are projected to enhance study habits and minimise absenteeism can play a significant role in improving student achievement. Student engagement can be tracked with the help of learning analytics systems and timely feedback to help educators to determine at-risk students and apply specific support measures. The wider context of learning analytics emphasizes its possibility to change educational processes through personalized learning and the use of data to make decisions (Viberg et al., 2018). Moreover, the role of engagement that is traced in extracurricular activities in this study indicates that fostering the whole student development can help to achieve better academic results. In order to produce well-rounded students and increase student motivation, educational institutions ought to contemplate using academic and co-curricular programs.

The study adds to the existing body of research on learning analytics by illustrating how one can employ empirical analysis on a multidimensional dataset. The results are consistent with the studies that have been conducted on student profiling, and that recommend the need to be aware of the differences in learning behaviors and performance between different individuals

(ISSN: 2208-2115)
Vol. 12, Iss. 02, 2026
DOI:10.65327/es.v12i2.2536

(Tempelaar et al., 2018). The study offers insights by recognizing the most important predictors of academic achievement that may be used to develop a personalized learning strategy and adaptive educational systems. Additionally, the combination of behavioral, demographic, and contextual variables of this study fills the gaps in the past studies that tend to focus on single variables. The holistic methodology embraced in this case improves the knowledge of intricate relationships among variables and their effect on student outcomes.

The results of this paper are consistent with the general trends of the studies on EDM and LA. The last year of studying, as shown in recent surveys, has highlighted the increasing use of data-driven approaches to the study of student behavior and prediction of student performance (Romero & Ventura, 2020). In this regard, this current research validates such tendencies since it proves the significance of empirical analysis in order to identify the most important determinants of the students' performance. In this current analysis, it has been pointed out that the performance of the students in their studies is viewed as an aspect of a complex set of behavioral participation, attendance, and other contextual issues. The use of data analytics is very efficient in understanding such relationships.

5. Conclusion

Empirical analysis of an empirical study about the performance of students in academics based on their learning behaviors, engagement patterns, and socio-demographic characteristics. Other factors, such as the time devoted to studying, absenteeism, tutoring, and parental assistance, are vital in relation to the outcome of the study, as demonstrated in the findings. In particular, the presence and consistency in studying of the pupils significantly affect the performance of the pupils, along with other factors. The finding proves to be relevant and hence makes it necessary to adopt the multi-dimensional approach to the performance of the students owing to the combination of descriptive, correlation, and regression analysis. It is also noted in the study that data-driven education, whereby data can be used to provide information on improving learning outcomes, is vital. The implications of this research include encouraging schools to foster good learning habits, reduce absenteeism, and build support mechanisms within the schools to help students learn better. The study uses a cross-sectional research design, but further studies would benefit from longitudinal data analysis and prediction modeling. In conclusion, the research makes a contribution to the area of educational data mining by offering relevant data related to student underperformance.

References

1. Akçapınar, G., Altun, A., & Aşkar, P. (2019). Using learning analytics to develop early-warning system for at-risk students. International Journal of Educational Technology in Higher Education, 16(1), 40.

2. Albreiki, B., Zaki, N., & Alashwal, H. (2021). A systematic literature review of student'performance prediction using machine learning techniques. Education Sciences, 11(9),

(ISSN: 2208-2115)
Vol. 12, Iss. 02, 2026
DOI:10.65327/es.v12i2.2536

552.

3. Alyahyan, E., & Düştegör, D. (2020). Predicting academic success in higher education: literature review and best practices. International journal of educational technology in higher education, 17(1), 3.

4. Asif, R., Merceron, A., Ali, S. A., & Haider, N. G. (2017). Analyzing undergraduate students' performance using educational data mining. Computers & education, 113, 177-194.

5. Aydoğdu, Ş. (2020). Predicting student final performance using artificial neural networks in online learning environments. Education and Information Technologies, 25(3), 1913-1927.

6. Bakhshinategh, B., Zaiane, O. R., ElAtia, S., & Ipperciel, D. (2018). Educational data mining applications and tasks: A survey of the last 10 years. Education and Information Technologies, 23(1), 537-553.

7. Boonk, L., Gijselaers, H. J., Ritzen, H., & Brand-Gruwel, S. (2018). A review of the relationship between parental involvement indicators and academic achievement. Educational research review, 24, 10-30.

8. Broadbent, J., & Poon, W. L. (2015). Self-regulated learning strategies & academic achievement in online higher education learning environments: A systematic review. The internet and higher education, 27, 1-13.

9. Castro, M., Expósito-Casas, E., López-Martín, E., Lizasoain, L., Navarro-Asencio, E., & Gaviria, J. L. (2015). Parental involvement on student academic achievement: A meta-analysis. Educational research review, 14, 33-46.

10. Conijn, R., Snijders, C., Kleingeld, A., & Matzat, U. (2016). Predicting student performance from LMS data: A comparison of 17 blended courses using Moodle LMS. IEEE Transactions on Learning Technologies, 10(1), 17-29.

11. Conijn, R., Van den Beemt, A., & Cuijpers, P. (2018). Predicting student performance in a blended MOOC. Journal of Computer Assisted Learning, 34(5), 615-628.

12. Elkharoua, R. (2024). Students performance dataset. Kaggle. https://www.kaggle.com/datasets/rabieelkharoua/students-performance-dataset

13. Gubbels, J., Van der Put, C. E., & Assink, M. (2019). Risk factors for school absenteeism and dropout: A meta-analytic review. Journal of youth and adolescence, 48(9), 1637-1667.

14. Ifenthaler, D., & Yau, J. Y. K. (2020). Utilising learning analytics to support study success in higher education: a systematic review. Educational technology research and

development, 68(4), 1961-1990.

15. Jovanović, J., Gašević, D., Dawson, S., Pardo, A., & Mirriahi, N. (2017). Learning analytics to unveil learning strategies in a flipped classroom. The Internet and Higher Education, 33, 74-85.

16. Keppens, G. (2023). School absenteeism and academic achievement: Does the timing of the absence matter?. Learning and Instruction, 86, 101769.

17. Khan, A., & Ghosh, S. K. (2021). Student performance analysis and prediction in classroom learning: A review of educational data mining studies. Education and information technologies, 26(1), 205-240.

18. Matcha, W., Uzir, N. A. A., Gašević, D., & Pardo, A. (2019). A systematic review of empirical studies on learning analytics dashboards: A self-regulated learning perspective. IEEE transactions on learning technologies, 13(2), 226-245.

19. Romero, C., & Ventura, S. (2020). Educational data mining and learning analytics: An updated survey. Wiley interdisciplinary reviews: Data mining and knowledge discovery, 10(3), e1355.

20. Tempelaar, D., Rienties, B., Mittelmeier, J., & Nguyen, Q. (2018). Student profiling in a dispositional learning analytics application using formative assessment. Computers in Human Behavior, 78, 408-420.

21. Viberg, O., Hatakka, M., Bälter, O., & Mavroudi, A. (2018). The current landscape of learning analytics in higher education. Computers in human behavior, 89, 98-110.

22. Waheed, H., Hassan, S. U., Aljohani, N. R., Hardman, J., Alelyani, S., & Nawaz, R. (2020). Predicting academic performance of students from VLE big data using deep learning models. Computers in Human behavior, 104, 106189.