EPH-International Journal of Applied Science
Volume - 12, Issue - 02
April- 2026
DOI https://doi.org/10.69980/2qeq8n66

LEARNING ANALYTICS FOR STUDENT SUCCESS PREDICTION

Manoj Kumar K S

Assistant Professor, Departement of English, English Literature, St. Claret College, Autonomous Bengaluru,
Email ID: Manoj2492@gmail.com, ORCHID:0009-0001-0245-3340

Received:-09-03-2026, Revised:-12-04-2026, Accepted:-20-04-2026, Published:-27-04-2026

Abstract
Learning analytics enables higher education institutions to use demographic, academic, assessment, registration, and Virtual Learning Environment data to identify students who may need early support. This retrospective predictive modeling study analyzed 32,593 student–module–presentation enrollments representing 28,785 unique students from the Open University Learning Analytics Dataset. Student success was defined as Pass or Distinction, while Fail or Withdrawn outcomes were classified as unsuccessful. Only information available during the first 30 course days was used to reduce data leakage. Predictors included early assessment scores, assessment completion, Virtual Learning Environment clicks, active learning days, unique resources accessed, educational background, credits studied, and module characteristics. The data were split into 26,122 training and 6,471 independent test observations using student-grouped partitioning. Logistic regression and random forest models were evaluated using accuracy, balanced accuracy, sensitivity, specificity, precision, F1-score, receiver operating characteristic area under the curve, precision-recall area under the curve, and Brier score. The random forest achieved the best performance, with 76.03% accuracy and a receiver operating characteristic area under the curve of 0.843, while logistic regression produced comparable results. Academic module, early assessment score, educational qualification, active learning days, and total platform activity were the leading predictors. These findings support transparent, ethically governed early-warning systems linked to timely, student-centered interventions in online higher education.

Keywords: *Academic performance, Early-warning systems, Learning analytics, Machine learning, Virtual Learning Environment*

Introduction

Digital platforms are used in higher education institutions that allow the gathering of information about student engagement, assessment work, access to courses, and students' learning behavior. These data have provided opportunities to explore students' experiences of the online learning environment and the relationship between those students' experiences and academic achievement. Learning analytics is the process of gathering, quantifying, analysing, and interpreting educational data in order to enhance the learning and teaching process and the decisions made by educational institutions. When used, it can help to spot students that might need assistance before it becomes irreversible that they are not performing well, withdrawing from school, or failing to succeed academically. Studies have shown that learning analytics can impact learning success when analytical results are applied to actionable, student-centered learning interventions in a timely fashion, as opposed to only reporting on the experiences of students in the past (Ifenthaler & Yau, 2020).

The value of learning analytics is directly tied to learning design because there are multiple levels of meaning in students' activity, which correspond to the design of the course, assessment, resource, and interactions. The total number of clicks, submissions, and accesses is not meaningful without context of the learning design. Data can be used in conjunction with course design to understand if students are achieving desired progress, and if learning activities are driving meaningful outcomes (Blumenstein, 2020). In addition, analytics can have a positive impact on the teaching process by providing information about the gaps in student participation, allowing teachers to adapt their teaching methods and engage students through evidence based on their learning behavior (Naujokaitienė et al., 2020).

The development of student success prediction is one of the significant applications of educational data mining. Predictive models are based on demographic, academic, behavioural, and assessment variables, which are used for estimating the probability of outcomes. Possible outcomes are pass, distinction, failure, withdrawal, and completion. Classification approaches like logistic regression, decision trees, random forest, support vector machine, and neural network are widely used in predicting academic performance in the higher education system, as seen from the systematic evidence (Alwarthan et al., 2022). Engagement variables are critical as they are indicative of the consistency with which students engage in learning activities, connect to learning resources, and complete learning assessments. Active days, frequency of interaction, resource use, and submission behaviour have been observed to correlate with academic performance multiple times in the past (Johar et al., 2023).

The use of technological developments has broadened learning analytics from traditional statistical modeling. Artificial intelligence, automated feedback, and large language models could help to create more adaptive and personalised education systems; however, it is crucial to ensure transparency, validation, and governance in their use (Mazzullo et al., 2023). Predictive analytics has particular significance in the realm of distance learning, where there may be little face-to-face interaction, making it hard to identify when students are becoming disengaged. Students at risk of dropping out can be identified early and receive targeted support for contact, tutoring, study guidance, etc. to prevent dropouts (Queiroga et al., 2020). Ensemble models have also proven to be useful in determining students' risk of academic failure in online learning systems, especially when several behavioural and academic factors are combined (Karalar et al., 2021).

Prediction is not just about classification; it's about taking action. Monitoring systems using artificial intelligence can be instrumental in institutions detecting poor performance and taking proactive measures before final results are known (Khan et al., 2021). Thus, early prediction is more useful than end-of-course prediction as it affords adequate time for intervention. The online learning activities in the early part of a course can provide predictive information, leading to earlier identification of academic risk (Wen & Juan, 2023). Prior education and previous attempts, as well as study load and other context factors, are also relevant, as these factors may impact preparedness and persistence (Köhler et al., 2023).

Although evidence has been increasing, prediction remains difficult due to several interacting factors which influence student success. Demographic-focused models could fail to capture the concept of changing learning behaviours, as could platform activity-focused models to the extent that they do not include educational background and assessment performance. The choice of a more comprehensive approach should include demographic, academic, registration and assessment data, as well as a Virtual Learning Environment approach, and should not allow for a leakage of information after the prediction point. Machine learning techniques can be helpful in facilitating this integration and identifying intricate patterns linked to successful and unsuccessful results (Guanin-Fajardo et al., 2024). For this reason, this study utilized learning analytics to predict students' success based on early data from the course, compared the performance between logistic regression and random forest, and found the most influential variables to predict students' success.

Objectives of the study:

1. To compare early assessment and Virtual Learning Environment engagement between successful and unsuccessful students.
2. To develop and evaluate logistic regression and random forest models for predicting student success.
3. To identify the demographic, academic, assessment, and engagement variables contributing most strongly to prediction.

2. Methodology

2.1 Study Design, Data Source, and Population

A secondary data analysis and predictive modelling study was conducted, which is a retrospective study using the Open University Learning Analytics Dataset (Kuzilek et al., 2015). Demographic, registration, assessment, module, and Virtual Learning Environment data were anonymized and included in the dataset of students enrolled in the distance-learning modules. The enrolment in the student–module–presentation unit was analysed. Data were integrated and validated, and all enrolments with a recorded final academic outcome were included. Students may have made more than one observation if they were enrolled in more than one module/presentation. The student, module, presentation, assessment and learning-resource identifiers were used to merge the student information, registration, assessment, course and Virtual Learning Environment files. Assessment and engagement records were combined to create one 'analytical' record per enrollee. A final dataset of 32,593 enrolments of 28,785 unique students was obtained.

2.2 Outcome Definition and Prediction Window

The main outcome was student success, which is determined by the resulting final academic outcome. Students who did not obtain a Pass or Distinction final result were deemed unsuccessful, and those who obtained a final result of Fail or Withdrawn were also considered unsuccessful. For descriptions, the original four outcome categories were kept, and for the predictive modelling, successful enrolment was considered as the positive class. The early prediction window of 30 days was used. This only included those that have been recorded in the assessment submission and the Virtual Learning Environment from course day 0 to day 30. Data collected after this time were omitted (later assessment scores, further engagement, withdrawal dates after the prediction time, etc.) to minimize data leakage.

2.3 Predictor Variables and Feature Engineering

Predictors were clustered under the headings of demographic, academic, registration, assessment, and engagement. Demographic variables were age band, Index of Multiple Deprivation band, disability status, highest qualification attained, region and gender. Academic variables were: module, module presentation, number of previous attempts, studied credits and the length of the presentation. The following were used as registration variables: Registration lead time and Registration status before commencement of the module. The number of submitted assessments, mean early score, weighted early score, proportion of assessments completed before the due date, late submissions, and missing early assessment scores were all considered early assessment variables. The engagement variables were total number of clicks in the Virtual Learning Environment, number of days spent learning, number of unique resources used, mean number of clicks per day that were spent learning, and inactive days. If the categorical values were missing, they were represented as their own unknown category; if the continuous values were missing, they were filled in with the median value from the training data. When there was no digital or recorded virtual learning environment activity, it was considered 0 engagement. All imputation, encoding, scaling, and feature generation were fitted on the training set and applied unchanged to the test set.

2.4 Descriptive and Comparative Analysis
The categorical variables were described in terms of frequencies and percentages. The summary statistics were medians and interquartile ranges, as these were the main assessment and engagement variables which were not normally distributed. Mann–Whitney U test was used to compare the successful and unsuccessful students in the continuous variables. The significance of statistical values was calculated with a 2-sided p-value less than 0.05.

2.5 Data Partitioning and Predictive Model Development
The analytical sample was split into a training sample of about 80% and a test sample of independent data of 20%. To avoid enrollment data from the same student being split between the training and test data sets, the enrollment data was partitioned according to groups. The training data set had 26122 enrolments and the independent test data set had 6471 enrolments. Two models (logistic regression and random forest) were built. The baseline model was logistic regression, and the nonlinear relationship and interactions among predictors were evaluated using random forest. A moderate imbalance in successful and unsuccessful enrolments was remedied by the use of class weighting. Logistic regression was carried out after encoding categorical variables and standardization of continuous variables. The training dataset was used to fit the models, with no changes, and the independent test dataset was used to evaluate the models.

2.6 Model Evaluation, Predictor Importance, and Software
The accuracy, balanced accuracy, sensitivity, specificity, precision, F1-score, receiver operating characteristic area under the curve, precision-recall area under the curve, and Brier score were used to measure model performance. Sensitivity was the number of successful enrolments who were correctly identified, and specificity was the number of unsuccessful enrolments who were correctly identified. A probability threshold of 0.50 was used for the binary classification, and a confusion matrix was constructed to provide a report of correct and incorrect classifications. To gauge the importance of each predictor in the random forest model, permutation importance was computed on the independent test set by randomly permuting each predictor in the model, which assesses the loss of predictive power due to permuting the predictor. Python was used for data integration, preprocessing, statistical analysis, predictive modelling, model evaluation, and visualization.

3. Results
3.1 Study Sample and Distribution of Academic Outcomes
There were 28,785 individual students enrolled in 32,593 enrolments in student–module–presentation combinations that made up the final analytical data set. Pass was the most frequent individual academic outcome, recorded in 12,361 (37.93%) enrolments, followed by withdrawal in 10,156 (31.16%), failure in 7,052 (21.64%), and distinction in 3,024 (9.28%). Following binary classification, there were 15,385 (47.20%) successful enrolments and 17,208 (52.80%) unsuccessful enrolments. The final academic outcomes are displayed in Table 1. The graph shown in Figure 1 shows the distribution of student enrolments by final academic outcomes.

Table 1. Distribution of final academic outcomes

Academic outcome	N	%
Distinction	3,024	9.28
Pass	12,361	37.93
Fail	7,052	21.64
Withdrawn	10,156	31.16
Successful	15,385	47.20
Unsuccessful	17,208	52.80
Total	32,593	100.00

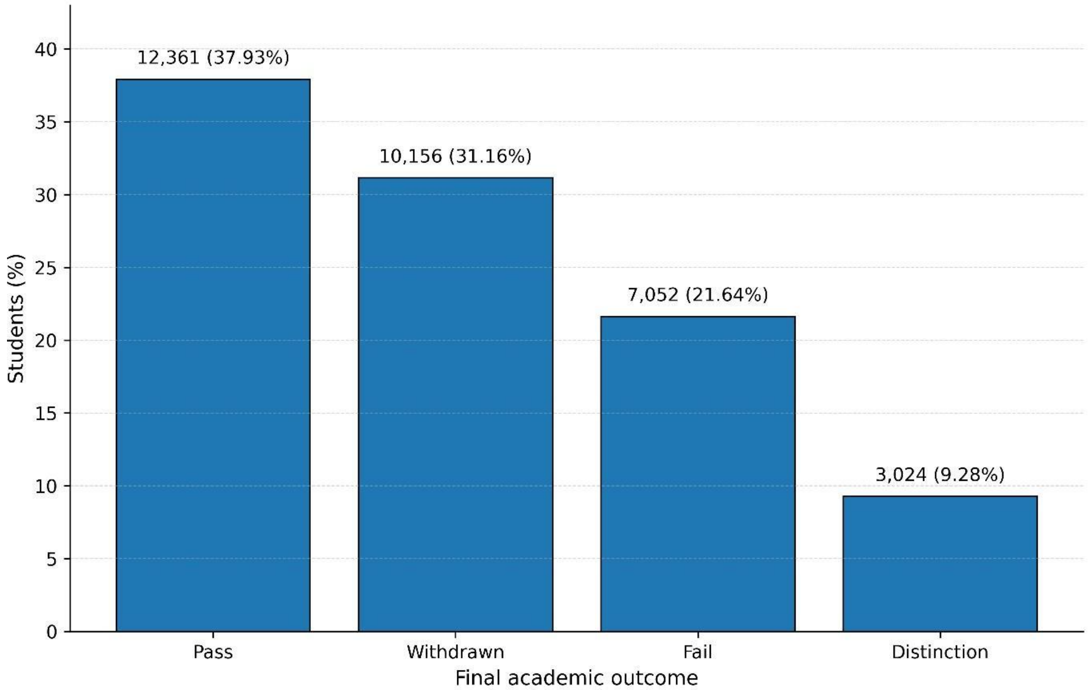

Figure 1. Distribution of Final Academic Outcomes Among Student Enrolments

3.2 Early Virtual Learning Environment Engagement

Students who were successful engaged significantly more with their Virtual Learning Environment during the first 30 days of learning. Among the successful students, the median number of total clicks on the VLE was 253, and among unsuccessful students, it was 71; the median number of active learning days was 13 for successful students and five for unsuccessful students. Successful students also used a median of 23 different VLE resources, whereas students who were unsuccessful used a median of 12 different VLE resources. All differences between groups were statistically significant ($p<0.001$). Table 2 shows early VLE engagement by student-success status.

Table 2. Early Virtual Learning Environment engagement according to student-success status

Variable	Successful median	Successful IQR	Unsuccessful median	Unsuccessful IQR	p-value
Total VLE clicks	253	381	71	217	<0.001
Active learning days	13	13	5	11	<0.001
Unique VLE resources	23	25	12	24	<0.001

3.3 Early Assessment and Academic Characteristics

There were significant differences in early assessment performance and completion between successful and unsuccessful students. The median scores for early assessment for those who succeeded and those who failed were 80.00 and 70.00, respectively, and the median weighted scores for those who succeeded and failed were 3.75 and 0.00, respectively. The median proportion of students who completed the assessment was 1.00 for successful students and 0.50 for unsuccessful students. There was also a statistically significant difference between the numbers of students who had attempted to register before and the number who had studied credits before, with the median number of previous attempts being zero for both groups, and between the length of time before students registered and the number of credits they had studied before, with the median number of credits registered being 25, despite there being no statistically significant difference between the two groups. Table 3 shows early assessment and academic characteristics by student-success status.

Table 3. Early assessment and academic characteristics according to student-success status

Variable	Successful median	Successful IQR	Unsuccessful median	Unsuccessful IQR	p-value
Mean early assessment score	80.00	19.00	70.00	25.00	<0.001
Weighted early assessment score	3.75	8.50	0.00	4.00	<0.001
Assessment completion proportion	1.00	0.00	0.50	1.00	<0.001
Previous attempts	0.00	0.00	0.00	0.00	<0.001
Studied credits	60.00	30.00	60.00	60.00	<0.001
Registration lead time, days	53.00	66.00	59.00	77.00	<0.001

3.4 Predictive Model Performance

A student-grouped partition yielded 26,122 training and 6,471 independent test observations. Logistic regression had an accuracy of 75.69%, a balanced accuracy of 76.07%, a sensitivity of 83.53%, a specificity of 68.60%, and the receiver operating characteristic (ROC) area under the curve of 0.837. The overall performance of the random forest was slightly better, with an accuracy of 76.03%, a balanced accuracy of 76.36%, specificity of 69.89%, precision of 71.33%, F1 score of 76.65%, receiver operating characteristic area under the curve of 0.843, and precision-recall area under the curve of 0.795. The Brier score was also slightly lower at 0.160. Logistic regression and random forest models were compared based on the performance shown in Figure 2. The comparative performance of the prediction models is presented in Table 4.

Table 4. Performance of the student-success prediction models

Metric	Logistic regression	Random forest
Accuracy, %	75.69	76.03
Balanced accuracy, %	76.07	76.36
Sensitivity, %	83.53	82.82
Specificity, %	68.60	69.89
Precision, %	70.64	71.33
F1-score, %	76.55	76.65
ROC AUC	0.837	0.843
PR AUC	0.784	0.795
Brier score	0.162	0.160

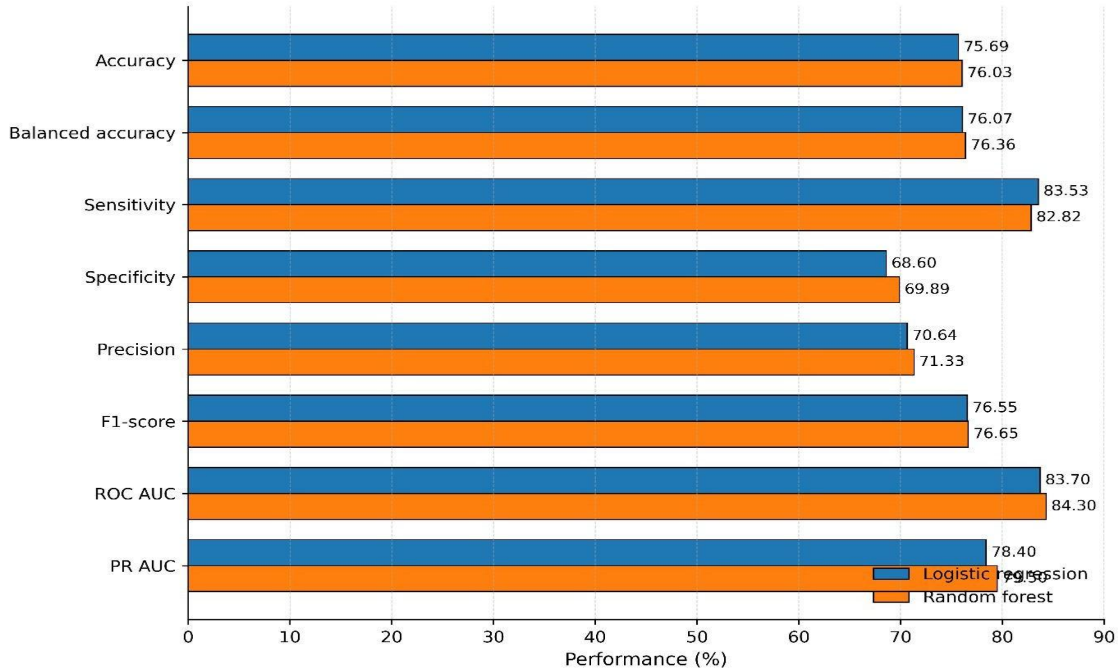

Figure 2. Comparative Performance of Logistic Regression and Random Forest Models for Student-Success Prediction

3.5 Classification Performance of the Random Forest Model

In the independent test sample, the random forest had an accuracy of 2545/2545 and 2375/2375 for successful and unsuccessful enrolments, respectively. It misidentified 528 successful enrolments as unsuccessful, and 1,023 unsuccessful enrolments as successful. These results were similar to a sensitivity of 82.82% and specificity of 69.89%, which showed a better identification of successful students than unsuccessful students.

3.6 Relative Importance of Predictors

The two most influential variables in the random forest model were academic module (Permutation Importance 0.0476) and mean assessment score within 30 days (Permutation Importance 0.0236). Other factors that helped predict were highest educational qualification, active VLE days, total number of VLE clicks, proportion of assessments completed, total number of inactive days, weighted early assessment score, number of credits studied, and absence of an early assessment score. The results suggested that student success classification was supported by multiple factors: early academic performance, sustained digital engagement, background factors, and factors related to the module. The relative importance of the predictors in the random forest model is shown in Figure 3. The most important predictors in the random forest model are displayed in Table 5.

Table 5. Leading predictors in the random forest model

Rank	Predictor	Permutation importance
1	Academic module	0.0476
2	Mean assessment score within 30 days	0.0236
3	Highest educational qualification	0.0125
4	Active VLE days	0.0106
5	Total VLE clicks	0.0085
6	Assessment completion proportion	0.0069
7	Inactive days	0.0065
8	Weighted early assessment score	0.0048
9	Studied credits	0.0044
10	Missing early assessment score	0.0042

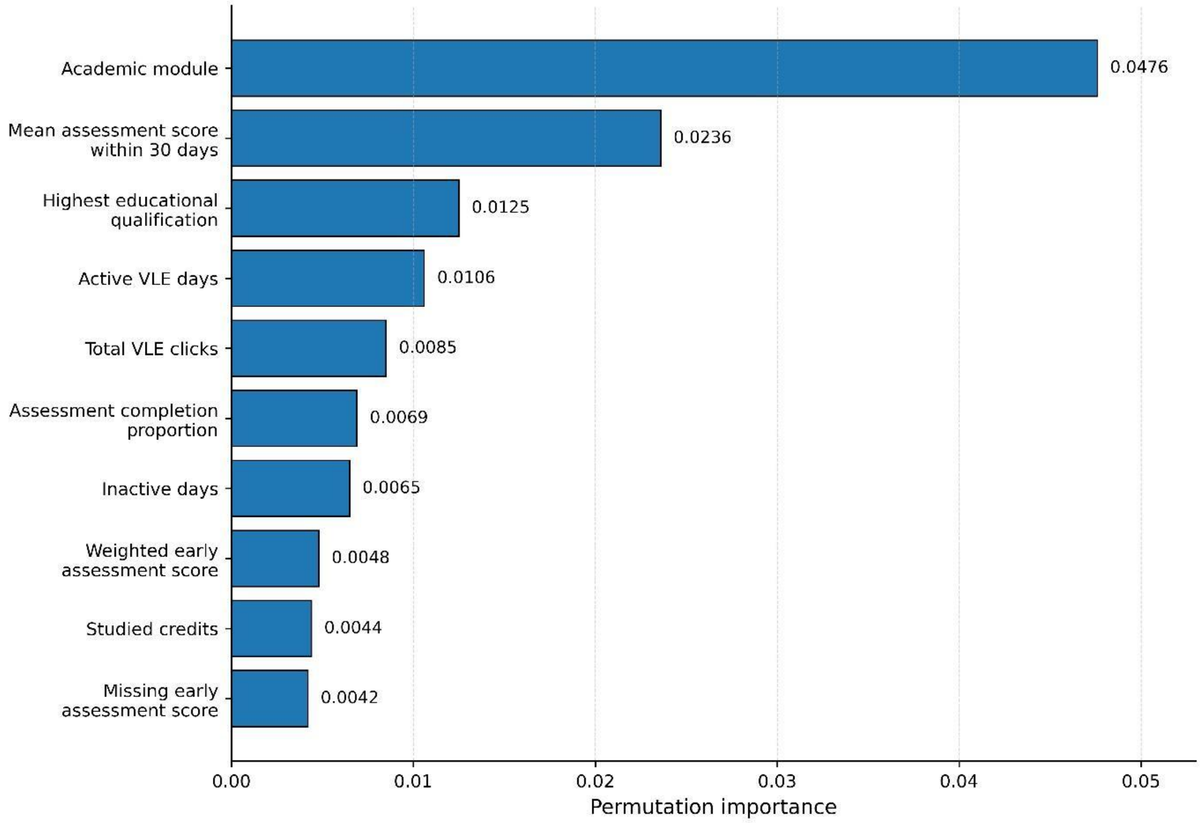

Figure 3. Relative Importance of Predictors in the Random Forest Student-Success Model

4. Discussion

In this study, the researchers were able to successfully discriminate in the first 30 days whether students would be successful or not with good accuracy. The random forest model was slightly better, but the difference was not large, compared to logistic regression. The significant predictive contributors were the early assessment score, assessment completion, days spent in the Virtual Learning Environment, total days spent in the platform, educational background, and academic module. The results align with the idea that student-related features, in combination with the instructional context, impact academic outcomes, where some indicators of student behavior are particularly salient for predicting outcomes (Jovanović et al., 2021).

The attitude of early engagement was associated with later success. Students who did well charted a higher number of clicks, more active learning days, and a wider variety of resources. The additional value of the regular participation of students seemed to be more informative than the total number of clicks, with activity days being more informative than the total number of active learning days. However, just because someone has clicked on a link does not necessarily mean that they learned something during that click-through, as many times they might have navigated simply because they were confused, they had a technical issue, or because they were just clicking on links without really reading or understanding the material. Their interpretation is enhanced with assessment behaviour and temporal aspects of participation. It has been demonstrated that if carefully chosen and properly combined and analysed, behavioural, cognitive, and contextual attributes can lead to better performance prediction when combined with multimodal learning data (Chango et al., 2021).

Performance on the initial assessment was also a key component of prediction. Successful students had higher scores, and more of them finished assessments, and were less likely to have an early score. Incomplete or missing assessment activity can be an indicator of disengagement, delayed participation, academic challenges, or barriers to participation. Institutions can respond by reminding, tutor contact, study-planning support, and/or technical assistance. Data must differentiate students in need of support from students who are missing data due to administrative issues. Explanations are important, as false reassurance can result in non-supported and vulnerable students, and false risk labels can result in unnecessary support (Hlosta et al., 2022).

Random forest has the best accuracy, specificity, F1 score, receiver operating characteristic area under the curve, and precision-recall area under the curve. The sensitivity was slightly higher, and the discrimination was almost as high with logistic regression. So the complexity of the model didn't provide much additional benefit. When the institutions need to interpret their predictions for educators, administrators, and students, simpler and interpretable models may be more desirable. There is evidence that interpretable models can achieve the same level of accuracy as their non-interpretable counterparts while also being fair (Kung & Yu, 2020).
The highest score in the random forest model was for the academic module. This likely resulted from variations in courses and assessment procedures, course design, and student populations, in the course prerequisites, and in instructional support provided to students. This suggests that a model trained locally could acquire some of the course-specific patterns of its outcomes. There is a need for module-level validation before implementation, especially when predictions are passed on to new subjects or academic presentations. Machine learning research also reveals that the accuracy of algorithms is dependent on the quality, relevance, and context of the variables fed into the algorithm (Ahmed, 2024).
The results are valid for the use of early-warning systems in online teaching. Assessment completion, student activity, first-month performance, and activity durations could be used as indicators to determine which students struggle to complete assessments, exhibit limited activity, exhibit extended periods of inactivity, and have poor initial performance. Dashboards and targeted feedback can support staff to make predictions and provide timely support, but whilst visual analytics can be useful, it is only if it is coupled with clear actions and teaching follow-up that it will lead to improved achievement (Alam et al., 2023). Learning analytics must work alongside academic judgement, not instead of it – especially if there are contextual factors that are not present. Consideration of ethics and equity is still a concern. Consent to Analytics might not be uniform across students, and consent patterns might be different across various groups of students, resulting in biased datasets and unequal representation (Li et al., 2022). Restricted access, privacy protection, explanation, and fairness audits are needed before use in operation. Limitations were a retrospective study design, the fact that data came from a single institution, only click-based engagement measures were used, and that personal circumstances (e.g., employment, motivation, Internet access, mental health) were not accounted for. While engagement is certainly a factor with academic performance in online learning, it is not a good indicator of student effort or ability alone (Sahni, 2023). These models need to be externally validated and evaluated in a prospective way before a high-stakes decision is made.

5. Conclusion

Data gathered from the first 30 days of study may perform well in predicting students' success, with good discriminatory power. The most important predictors of the assessment scores were early assessment scores, completion of the assessments, active use of the Virtual Learning Environment, total platform activity, level of education, and academic module. Overall, the random forest model performed slightly better than logistic regression, but both models gave comparable results, suggesting that simpler models could be appropriate for use in institutions. The results have relevance to the creation of early-warning systems that alert educators to students who might need some quick academic, technical, or motivational assistance. The use of such systems should help inform, not supplant, academic judgement, especially when there are factors outside the data set that impact student performance. The study had several limitations regarding generalizability and causal interpretation, including a retrospective design, limited to a single institution and solely using data from the platform. Requirements before implementation are: External validation, module-specific evaluation, fairness assessment, and prospective testing. Learning analytics can bolster student-support strategies in those cases where the predictions are clear, ethical, and connected to targeted interventions to prevent disengagement and failure or withdrawal from the learning environment in all manner of higher education contexts.

References

1. Ahmed, E. (2024). Student performance prediction using machine learning algorithms. Applied computational intelligence and soft computing, 2024(1), 4067721.
2. Alam, M. I., Malone, L., Nadolny, L., Brown, M., & Cervato, C. (2023). Investigating the impact of a gamified learning analytics dashboard: Student experiences and academic achievement. Journal of Computer Assisted Learning, 39(5), 1436-1449.
3. Alwarthan, S. A., Aslam, N., & Khan, I. U. (2022). Predicting student academic performance at higher education using data mining: A systematic review. Applied Computational Intelligence and Soft Computing, 2022(1), 8924028.
4. Blumenstein, M. (2020). Synergies of learning analytics and learning design: A systematic review of student outcomes. Journal of Learning Analytics, 7(3), 13-32.
5. Chango, W., Cerezo, R., Sanchez-Santillan, M., Azevedo, R., & Romero, C. (2021). Improving prediction of students' performance in intelligent tutoring systems using attribute selection and ensembles of different multimodal data sources. Journal of Computing in Higher Education, 33(3), 614-634.
6. Guanin-Fajardo, J. H., Guaña-Moya, J., & Casillas, J. (2024). Predicting academic success of college students using machine learning techniques. Data, 9(4), 60.
7. Hlosta, M., Herodotou, C., Papathoma, T., Gillespie, A., & Bergamin, P. (2022). Predictive learning analytics in online education: A deeper understanding through explaining algorithmic errors. Computers and Education: Artificial Intelligence, 3, 100108.
8. 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.
9. Johar, N. A., Kew, S. N., Tasir, Z., & Koh, E. (2023). Learning analytics on student engagement to enhance students' learning performance: A systematic review. Sustainability, 15(10), 7849.
10. Jovanović, J., Saqr, M., Joksimović, S., & Gašević, D. (2021). Students matter the most in learning analytics: The effects of internal and instructional conditions in predicting academic success. Computers & Education, 172, 104251.
11. Karalar, H., Kapucu, C., & Gürüler, H. (2021). Predicting students at risk of academic failure using ensemble model during pandemic in a distance learning system. International Journal of Educational Technology in Higher Education, 18(1), 63.
12. Khan, I., Ahmad, A. R., Jabeur, N., & Mahdi, M. N. (2021). An artificial intelligence approach to monitor student performance and devise preventive measures. Smart Learning Environments, 8(1), 17.
13. Köhler, J., Hidalgo, L., & Jara, J. L. (2023). Predicting students' outcome in an introductory programming course: leveraging the student background. Applied Sciences, 13(21), 11994.
14. Kung, C., & Yu, R. (2020, August). Interpretable models do not compromise accuracy or fairness in predicting college success. In Proceedings of the seventh acm conference on learning@ scale (pp. 413-416).
15. Kuzilek, J., Hlosta, M., & Zdrahal, Z. (2015). Open University Learning Analytics dataset [Dataset]. UCI Machine Learning Repository. https://archive.ics.uci.edu/dataset/349/open%2Buniversity%2Blearning%2Banalytics%2Bdataset
16. Li, W., Sun, K., Schaub, F., & Brooks, C. (2022). Disparities in students' propensity to consent to learning analytics. International Journal of Artificial Intelligence in Education, 32(3), 564-608.
17. Mazzullo, E., Bulut, O., Wongvorachan, T., & Tan, B. (2023). Learning analytics in the era of large language models. Analytics, 2(4), 877-898.
18. Naujokaitienė, J., Tamoliūnė, G., Volungevičienė, A., & Duart, J. M. (2020). Using learning analytics to engage students: Improving teaching practices through informed interactions. Journal of New Approaches in Educational Research, 9(2), 231-244.
19. Queiroga, E. M., Lopes, J. L., Kappel, K., Aguiar, M., Araújo, R. M., Munoz, R., ... & Cechinel, C. (2020). A learning analytics approach to identify students at risk of dropout: A case study with a technical distance education course. Applied Sciences, 10(11), 3998.

20. Sahni, J. (2023). Is learning analytics the future of online education?: assessing student engagement and academic performance in the online learning environment. International Journal of Emerging Technologies in Learning (iJET), 18(2), 33-49.
21. Wen, X., & Juan, H. (2023). Early prediction of students' performance using a deep neural network based on online learning activity sequence. Applied Sciences, 13(15), 8933.