International Journal for Research in Educational Studies
(ISSN: 2208-2115)
Vol. 12, Iss. 02, 2026
DOI:10.65327/es.v12i2.2535

Digital Learning Engagement, Learner Diversity, and Academic Achievement in Secondary Education: A Quantitative Analysis of LMS Interaction Data.

Dr. Rabia Al-Mur Al-Dhuhli

Associate Professor of Educational Administration, Department of Education and Cultural Studies, University of Nizwa, Nizwa, Sultanate of Oman,Email: rabeealthuhli@unizwa.edu.om, Orchid Id: https://orcid.org/0000-0002-5949-5912

Received: 02/06/2026 Revised: 05/07/2026 Acceptance: 13/07/2026 Published: 21/07/2026

ABSTRACT

Digital learning contexts produce behavioural data that could inform more accurate understanding of students' learning and engagement, but often these behavioural data are studied in isolation with less consideration of learner diversity and context. With a cross sectional, explanatory-predictive design, this study examined the relationship among digital learning engagement, learner characteristics, and academic achievement in secondary education. A total of 281 middle and high school records from the xAPI-Edu-Data dataset were analysed with engagement defined as the number of times students raised their hand, visited a resource, viewed an announcement, and participated in discussion. The engagement index was standardised and tested for validity using non-parametric tests and ordinal logistic regression, and compared with multinomial logistic regression, random forest and gradient boosting models by repeated stratified five times cross validation. There were positive and consistent associations for digital engagement with the low, middle and high achievement groups, and a strong positive association with ordered achievement. The odds of higher achievement levels were about 6.7 times greater for a one-SD increase in engagement in the adjusted ordinal model. The best predictive model was the combination of random forest models, which had a macro area under the curve of 0.928 and an accuracy of about 79.5%. The findings show that combining LMS engagement and attendance, parent and learner-context factors offers a stronger foundation for achievement classification and data-informed student support than behavioural factors alone.

Keywords – digital learning engagement, learning management systems, academic achievement, learner diversity, educational data mining, secondary education

---

1. Introduction

The digital transition has completely changed the way teaching, learning, assessment and supporting students happens in secondary education. Learning management systems are intermediaries between student and teacher, their content and activities and the learner, and leave traces of these interactions that can be analyzed together with learning outcomes. The construct of digital learning engagement is therefore at the heart because it is observable when students engage with technology supported environments and can be used to identify those who could benefit from support. Research in high schools shows that digital education can improve student participation by providing a supportive learning environment that facilitates ongoing interaction and involvement (Aldhafeeri & Alotaibi, 2022). However, engagement is not always positive or consistent; sometimes students switch between engagement and disengagement as a result of the design of the tasks, their support, their motivation, and their learning environment (Bergdahl, 2022).

Digital engagement is not just about how they use the platform, it's about how it can help them in education. Embedding technology in pedagogy is a crucial component in effective integration, as it is not enough to have access for purposeful participation and competence development. Recent studies indicate that better integration is correlated with students' behavioral engagement and digital skills, highlighting the importance of identifying meaningful interactions from activity counts (Consoli et al., 2025). Meanwhile, the correlation between engagement and achievement is not constant and equal across learners, or over time. However, the findings from the current study align with previous longitudinal studies, which have found that these connections are contingent upon the time of the study, the characteristics of the student sample, and the starting point of the students' academic experience, meaning that indicators of engagement need to be understood in the context of differential, not universal, educational experience (Saqr et al., 2023).

The complexities make a research problem. While learning platforms do provide student interaction data, many analyses are descriptive, look at individual interactions or only use estimation data. This type of strategies offer little information as to whether the engagement measures explain variance beyond the variance for learner characteristics, attendance, parental involvement, and instruction context. They also provide inadequate information on whether statistical correlations correspond to good out-of-sample classification of academic achievement. The potential of using data-driven precision teaching to promote differentiated intervention in secondary schools is demonstrated by the application of educational data and analytic models (Wang et al., 2024). But where it comes to precision-oriented applications, it is crucial to have strong validation, transparent feature building, and clear distinction from predictive uses and causal statements.

Therefore, the present study considers digital engagement as a multidimensional behavior that is considered within the learner and context conditions. Related types of engagement include resource visits, announcement viewing, participation in discussion, and raised hand activity, and demographic, educational stage, subject and attendance, and parental variables are used to provide contextual information that could affect achievement patterns. LAR studies have demonstrated that analytic feedback is related to students' psychological needs and satisfaction in BL, which further highlights the need to use behavioral data for learning support and not for compliance monitoring (Ameloot et al., 2024). Likewise, homeworks systems for personalization through formative learning analytics show that the data can be used to design individualized responses to teaching, though the quality and relevance of the indicators that feed into the system determines the effectiveness of the solution (Rodríguez-Martínez et al., 2023).

The following empirical gap arises from the fact that the integration of construct validation, inferential modelling, comparative machine learning and class-sensitive evaluation were limited to one study in the field of secondary education. The associations between engagement and achievement, or between predictive performance and engagement have been explored in existing work, typically in isolation. There is limited evidence that a validated composite engagement construct is valid in an ordinal explanatory model and that the engagement-only specification, context-only specification, and combined specification are equally valid models that can be compared using repeated cross-validation. The consequence of this gap is that a model can have statistically significant coefficients, but not perform well in prediction, while another model can perform well in prediction, but not be well understood. To achieve a state of theoretical clarity, methodological rigor, and relevance in educational research, both requirements must be satisfied.

The study is answered with the xAPI-Edu-Data data set, limited to 281 middle and high school observations. Four behavioral engagement indicators, demographics and education data, parental participation and satisfaction data, attendance categories, and an ordinal achievement outcome (low, middle, and high) were included in the dataset. Its structure allows investigation of the role of digital engagement as an explanatory, and predictive, variable in addition to the learner-context variables. The first goal is to see if there are systematic differences in digital learning engagement by achievement categories and positive relationships between ordered achievement levels and digital learning engagement. The second aim is to estimate the independent contribution of digital engagement given the control of learner-diversity, educational and parental characteristics and attendance-related characteristics. The third goal is to evaluate the various models (engagement-only, context-only, and combined statistical and machine-learning models) for achieving the most reliable and consistent classification of children's achievement.

The research design used in this study is a cross sectional research design with an observational-exploratory-predictive approach. A standardized engagement index is developed from the raised hand participation, visits in the resources and viewing the announcements, and the amount of discussion or debate. This index is tested for internal consistency before inferential analysis. Spearman correlation and Kruskal-Wallis tests examine ordered relationships and group differences; proportional-odds ordinal logistic regression examines associations with achievement, adjusting for other factors. The performance is compared among the three methods: multinomial logistic regression, random forest, and gradient boosting by repeated stratified 5-fold cross-validation based on the following measures: accuracy, balanced accuracy, macro F1, and macro area under the receiver operating characteristic curve. This design has the benefits of being statistically interpretable and computationally validated, and reduces the dependence on a single modelling paradigm.

The scope of this study is limited. The data are one secondary data set that is primarily middle school and has a small high school sub-group. Academic achievement is measured categorically (not in terms of scores or growth over time), and learner diversity is represented by existing demographic, educational, family and attendance data (not by disability, socioeconomic status, language proficiency or conditions of digital access). The observational structure does not allow for causal conclusions and external generalizability is to be taken with a grain of salt. However, the data can be used for a targeted inquiry of behavioral engagement and classification of achievement within clear constraints.

The study has three significance. This, theoretically, explains the relation of a multidimensional construct of behavioral engagement with achievement beyond contextual adjustment. It has a methodologically integrated approach of construct validation, ordinal modelling, reduced-model comparison, repeated cross-validation, confusion-matrix analysis, and permutation-based feature importance. In practice, it assesses the potential of data from learning platforms and their context to help identify low achieving learners. These uses should be left to the discretion of the teacher and teacher trust: teacher trust is a driver of the adoption of artificial-intelligence-supported educational technologies (Nazaretsky et al., 2022). They need to also include privacy protections since learners' perceptions of how data is collected, controlled, and used by institutions influence their concerns with learning analytics (Mutimukwe et al., 2022). The study provides a grounded framework for data-informed and ethical secondary student support, by integrating explanatory and predictive evidence.

2. Literature Review

The research literature increasingly focuses on how digital learning engagement is a product of the combination of the affordances of the technologies, characteristics of the learners, and the design of the teaching. The benefits of emerging technologies in terms of engagement in STEAM environments rely on a purposeful pedagogical integration, not technology itself; as they offer new opportunities for interaction, creation and multimodal participation (Maričić & Lavicza, 2024). This is important because engagement does not only refer to observable frequency of use on the platform, but it is also shaped by affective and motivational processes. Engagement in online learning has been found to be associated with achievement emotions, suggesting that traces of behaviour may have an underlying psychological component, which is not necessarily captured in records of the learning-management-system (LMS) (Bakır-Yalçın & Usluel, 2024). Other evidence found indicates that emotions, metacognition and online learning readiness are intertwined predictors of engagement, which poses a challenge to the models which maintain engagement as a unidimensional behavioral construct (Fidan & Koçak Usluel, 2024).

The second theme is about the analytical significance of LMS interaction data. Predictive studies indicate that platform behaviors can be modeled to predict future behaviors of student interaction, however, features' quality, granularity of the time, and the methods used for the validation of the models can affect the usefulness of these models (Alshammari, 2024). The research results that are related to the study also indicate that combinations of access to resources, communication, and participation provide more information than individual elements, but with results differing across courses and institutions (Cabı, 2025). In the context of early-prediction in game-based learning, distributed representations also demonstrate the ability to predict knowledge states prior to the test; again, the motivation of the models is to predict knowledge, rather than to explain it clearly (Emerson et al., 2023). Learning-analytics interventions that leverage real-time feedback have also shown to enhance collaborative knowledge processes and group performance, but intervention studies are approached differently than observational datasets, and as such cannot be directly translated to secondary LMS records (Zheng et al., 2022).

A third theme is an engagement design-achievement link. While gamified formative assessment has shown to boost participation and performance, leaderboard effects can vary based on competition, discipline, and learner profile, suggesting potential variability across education contexts (Cigdem et al., 2024). There are models that provide a more comprehensive assessment capability but many do not have a comparator to an interpretable baseline or a class sensitive metric (Lam et al., 2024). Educational data-mining studies also show that the performance of the demographic, behavioural, and academic variables are all helpful in predicting educational outcomes, although they are often evaluated using inconsistent datasets and pre-processing methods, making them difficult to compare and replicate (Sarker et al., 2024). The literature available on deep learning in virtual environments is systematic, and there is strong evidence of the predictive power of deep learning, but there are also limitations with respect to: the small number of samples, imbalance, opacity, and weak external validation (Alnasyan et al., 2024).

The relevant literature thus presents a gap, because few studies combine a validated engagement construct with contextual learner variables with ordinal explanatory modelling within the framework of a repeated cross-validated comparison of reduced and combined predictive models, in secondary education. The current study aims to fill this gap by focusing on 281 middle- and high-school observations, an engagement index based on four behavioral measures, ordinal logistic regression, and comparative machine-learning analysis. It compares engagement-only, context-only and combined models, and combines interpretability with predictive rigor in an explicit validation context, but constrains itself by the cross-sectional, categorical and context specific nature of the data set.

3. Methodology

3.1 Research Design and Analytical Framework

The research design used in this study was empirical, observational, cross sectional and computational to explore the relationship between digital learning engagement and characteristics of the learners as regards academic achievement at secondary school level. The design was a mixture of statistical modelling to explain the data and machine-learning algorithms to predict because the research problem needed to describe relationships between the variables and to evaluate the prediction accuracy. The analytical framework defined academic achievement as an ordinal outcome through three complementary domains: digital engagement, learner diversity and

International Journal for Research in Educational Studies
(ISSN: 2208-2115)
Vol. 12, Iss. 02, 2026
DOI:10.65327/es.v12i2.2535

educational context. In formal terms, learning level i was modeled as Yi=f(Ei,Di,Ci), with the engagement indicators (Ei), the demographic (Di) and educational (Di) factors, and the parental and attendance (Ci) factors. The absence of an intervention, random assignment, baseline measure, and longitudinal follow up precluded the interpretation of all estimates as causal effects.

3.2 Dataset, Sample Selection, and Data Quality

The analysis was conducted on the provided xAPI-Edu-Data CSV file, which is comprised of 480 observations and 17 variables representing students' characteristics, classroom participation, learning-resource use, family involvement, attendance, and achievement classification (Amrieh et al., 2016). The analysis was performed on the target population in each of the contexts included in the dataset: students in secondary-level education. Probabilistic sampling was not used, instead, a criterion-based census of all eligible records was performed. Records with MiddleSchool or HighSchool StageID were included, while 199 records were not included because they did not correspond to the focus of the title of secondary education. The resulting sample of data, an analytical sample, was 281 observations: 248 middle school and 33 high school observations. There were 62 low, 132 middle and 87 high cases. No missing values or exact replicates were present in the data screened in the filtered sample. All transformations were carried out in repeatable Python pipelines, category labels were standardized and types of variables were checked. No specific personal details were included within the dataset and thus the data was only analysed at an aggregate level and reporting was limited to just that level, with no attempts made to re-identify the data.

3.3 Variable Operationalisation

Academic achievement was operationalised using Class (low, middle and high). Digital learning engagement was symbolized by raisedhands, VisITedResources, AnnouncementsView, and Discussion (participation, access to resources, attention to instructional communications, and discussion activity). The indicators were converted into z-scores and the arithmetic mean of the z-scores was used to create the Digital Learning Engagement Index. Learner-diversity and contextual predictors consisted of gender, NationalITy, StageID, GradeID, SectionID, Topic, Semester, Relation, ParentAnsweringSurvey, ParentschoolSatisfaction, StudentAbsenceDays. PlaceofBirth was not included in the main multivariable models due to conceptual overlap with NationalITy and potential for double encoding. One-hot encoding and setting one category as the reference category was employed for categorical predictors. The composite index was used for the principal ordinal model while continuous engagement variables were used as individual variables for the machine-learning models.

3.4 Construct Validation and Preliminary Analysis

Internal consistency of the four engagement indicators was evaluated by calculating Cronbach's alpha and the corrected item-total relationships were analyzed to determine defensibility for aggregation. Descriptive analysis consisted of frequencies and percentages for categorical variables and means, standard deviations, medians and interquartile ranges for numerical variables. Cross-tabulations and group-specific engagement summaries were used to look at achievement-group patterns. Since the outcome was of an ordinal nature and normality could not be assumed, the monotonic association between the engagement index and the ordered achievements was quantified using Spearman's rank correlation. A Kruskal-Wallis test was used to determine if there were differences in the distributions of engagement among the achievement groups. For omnibus tests of significance a multiplicity adjustment was performed for the pairwise comparisons, where appropriate. Prior to model estimation, sparse categories, extreme values, class imbalance, and bivariate redundancy were also explored preliminarily.

3.5 Explanatory Statistical Modelling

An ordinal logistic regression model with a proportional odds assumption calculated the odds of being in increasingly higher achievement levels, given membership in each one. The engagement index was modelled as the standardised index, while the theoretically relevant demographic, educational, parental, satisfaction, semester, stage, and attendance covariates were included. The parameters were estimated using maximum likelihood and presented as log-odds coefficients, odds ratios, 95% confidence intervals and two sided p-values. The model adequacy was assessed using the likelihood-based comparisons, Akaike and Bayesian information criteria, and the McFadden’s pseudo-R-squared. The problem of multicollinearity was evaluated through variance inflation factors and correlation diagnostics and the potentially dominant observations were investigated through residual diagnostics and influence diagnostics. The proportional-odds assumption was explored by comparing thresholds and using sensitivity models. This assessment was repeated by re-estimating the models without sparse categorical levels in individual engagement indicators and compared with the specification without the highly influential contextual variables.This was done again by re-estimating the models with the individual engagement indicators only (not sparse categorical levels) and compared to the specification without the highly influential contextual variables.

3.6 Predictive Modelling and Validation

Three sets of features were tested: engagement only, diversity-context only and both engagement and diversity-context.Three models were compared for predictive analysis: multinomial logistic regression, random forest and gradient boosting with three sets of features: engagement only, diversity-context only, and both engagement and diversity-context. One-hot encoding and scaling (when necessary) were performed in each resampling fold separately to avoid information leakage during pre-processing. Models were tested with three repetitions of repeated stratified 5-fold cross-validation maintaining the category proportions in each fold of the cross-validation. Primary metrics involved: accuracy, balanced accuracy, macro-averaged F1 and macro area under receiver operating characteristic curve against one-versus-rest approach. The final model was chosen based on the consistent performance of the model across the metrics and not based on 100% accuracy. Confusion matrix and class-specific prediction precision and recall and F1 estimates were created using out-of-fold predictions. The contribution of each predictor to the predictive performance was quantified by permutation importance. Several alternative sets of features, algorithms, random seeds and class-sensitive metrics were compared using sensitivity analyses. The workflows were scripted and the fixed seeds used for all analyses for reproducibility.All analyses performed in Python with pandas, NumPy, SciPy, statsmodels and scikit-learn. Preprocessing, estimation and validation and reporting decisions were captured along with version-controlled code.

4. Results

4.1 Sample Characteristics and Achievement Distribution

The secondary-education filter retained 281 of the original 480 observations. The analytical sample comprised 248 middle-school records and 33 high-school records. No missing values or exact duplicate observations were identified after filtering. The sample contained 184 male and 97 female learners. Academic achievement was distributed across 62 low-, 132 middle-, and 87 high-achievement observations. Table 1 presents the principal characteristics of the analytical sample.

Table 1. Distribution of the Secondary-Education Sample

Characteristic	Category	n	%
Educational stage	Middle school	248	88.3
	High school	33	11.7
Gender	Male	184	65.5
	Female	97	34.5
Academic achievement	Low	62	22.1
	Middle	132	47.0
	High	87	31.0

Table 1 shows that middle-school observations constituted most of the analytical sample. The middle-achievement category was the largest outcome group, whereas the low-achievement

category was the smallest.

4.2 Digital-Engagement Construct and Group Patterns

The four behavioural indicators demonstrated satisfactory internal consistency, with Cronbach's alpha equal to .822. Their aggregation into the Digital Learning Engagement Index was therefore retained. Mean values for raised-hand participation, resource visits, announcement viewing,

(ISSN: 2208-2115)
Vol. 12, Iss. 02, 2026
DOI:10.65327/es.v12i2.2535

discussion activity, and the composite index increased across the ordered achievement categories. Table 2 reports the group-specific engagement statistics.

Table 2. Digital-Engagement Indicators by Academic-Achievement Category

Achievement category	n	Raised hands, M (SD)	Resources visited, M (SD)	Announcements viewed, M (SD)	Discussion, M (SD)	Engagement index, M (SD)
Low	62	16.56 (16.28)	18.00 (19.32)	17.08 (16.49)	27.11 (22.37)	−1.02 (0.42)
Middle	132	54.34 (25.80)	61.55 (25.71)	44.22 (22.83)	46.15 (25.96)	0.06 (0.61)
High	87	75.87 (20.21)	79.93 (17.47)	58.08 (24.50)	61.56 (25.07)	0.64 (0.51)

Table 2 shows that all four engagement indicators followed the same ordered pattern, with the lowest means in the low-achievement category and the highest means in the high-achievement category. The largest absolute differences between low- and high-achievement observations occurred for resource visits and raised-hand participation.

Spearman's rank correlation identified a strong positive association between the engagement index and ordered achievement level, $\rho=.710$, $p<.001$. The Kruskal–Wallis test also detected significant differences in engagement-index distributions across the three achievement categories, $H(2)=147.19$, $p<.001$. Holm-adjusted pairwise Mann–Whitney tests were significant for low versus middle achievement ($U=604.00$, adjusted $p<.001$), low versus high achievement ($U=99.00$, adjusted $p<.001$), and middle versus high achievement ($U=2640.00$, adjusted $p<.001$). As illustrated in Figure 1, the mean standardised engagement index increased monotonically across the achievement categories.

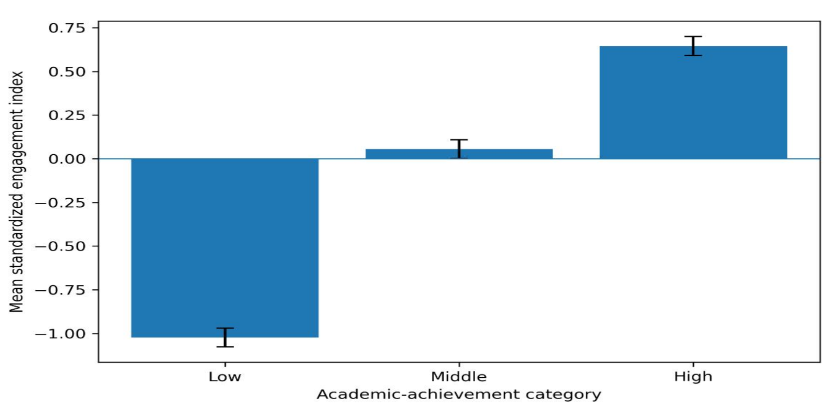

Figure 1. Mean Digital-Engagement Index across Achievement Categories

(ISSN: 2208-2115)
Vol. 12, Iss. 02, 2026
DOI:10.65327/es.v12i2.2535

Figure 1 shows that low-achievement observations were located substantially below the sample engagement mean, whereas high-achievement observations were located above it. The middle-achievement group was centred close to the overall sample mean.

4.3 Explanatory Model Results

The proportional-odds ordinal logistic model converged successfully using 281 observations. The estimated log likelihood was −144.40, with an Akaike information criterion of 308.79 and a Bayesian information criterion of 345.18. McFadden's pseudo-R2 was .511. Table 3 presents the estimated coefficients, standard errors, odds ratios, confidence intervals, and significance levels.

Table 3. Ordinal Logistic-Regression Estimates for Academic-Achievement Level

Predictor	β	SE	Odds ratio	95% CI for OR	p
Standardised engagement index	1.908	0.259	6.740	4.056–11.199	<.001
Female learner	1.120	0.333	3.065	1.596–5.884	.001
High-school stage	0.343	0.457	1.409	0.576–3.451	.453
Second semester	0.114	0.312	1.120	0.608–2.065	.715
Mother as responding parent	1.247	0.335	3.479	1.803–6.715	<.001
Parent answered survey	0.825	0.400	2.282	1.042–4.996	.039
Positive school satisfaction	0.086	0.395	1.090	0.502–2.366	.828
Fewer than seven absence days	2.905	0.440	18.269	7.717–43.247	<.001

Table 3 shows that the standardised engagement index was positively associated with placement in a higher achievement category after adjustment for the included learner and contextual characteristics. Female gender, maternal parental relationship, parental survey participation, and fewer than seven absence days also had statistically significant positive coefficients.

A one-standard-deviation increase in the engagement index corresponded to 6.74 times higher cumulative odds of membership in a higher achievement category, holding the remaining covariates constant. Learners with fewer than seven absence days had 18.27 times higher cumulative odds of higher achievement than learners with more than seven absence days. Female learners had 3.06 times higher cumulative odds than male learners, while records identifying the mother as the responding parent had 3.48 times higher cumulative odds than records identifying the father.

International Journal for Research in Educational Studies
(ISSN: 2208-2115)
Vol. 12, Iss. 02, 2026
DOI:10.65327/es.v12i2.2535

High-school stage, second-semester enrolment, and positive parental school satisfaction were not statistically significant in the adjusted model. Variance-inflation factors for the explanatory predictors ranged from 1.04 to 1.78, excluding the intercept, and did not indicate substantial predictor collinearity. Threshold-specific binary sensitivity models retained positive and statistically significant coefficients for engagement and low absence at both achievement thresholds.

4.4 Comparative Predictive Performance

Nine predictive specifications were evaluated by combining three algorithms with three feature sets. Performance was estimated through repeated stratified five-fold cross-validation with three repetitions. Table 4 reports the mean validation scores and their standard deviations across the 15 resampling folds.

Table 4. Repeated Cross-Validated Predictive Performance

Feature set	Algorithm	Accuracy, M (SD)	Balanced accuracy, M (SD)	Macro F1, M (SD)	Macro AUC, M (SD)
Engagement only	Multinomial logistic regression	.668 (.051)	.681 (.049)	.675 (.049)	.851 (.032)
Engagement only	Random forest	.723 (.058)	.726 (.062)	.726 (.058)	.875 (.041)
Engagement only	Gradient boosting	.706 (.058)	.713 (.059)	.713 (.056)	.855 (.047)
Diversity-context only	Multinomial logistic regression	.687 (.050)	.690 (.049)	.693 (.053)	.846 (.035)
Diversity-context only	Random forest	.696 (.045)	.685 (.046)	.695 (.048)	.851 (.032)
Diversity-context only	Gradient boosting	.718 (.045)	.713 (.047)	.718 (.048)	.861 (.036)
Combined	Multinomial logistic regression	.783 (.033)	.788 (.029)	.790 (.031)	.903 (.019)
Combined	Random forest	.795 (.033)	.791 (.037)	.799 (.033)	.928 (.017)
Combined	Gradient boosting	.785 (.038)	.784 (.050)	.790 (.040)	.911 (.023)

(ISSN: 2208-2115)
Vol. 12, Iss. 02, 2026
DOI:10.65327/es.v12i2.2535

Table 4 shows that each combined model exceeded its corresponding engagement-only and diversity-context-only specification in accuracy and macro AUC. The combined random forest produced the highest mean accuracy, macro F1, and macro AUC, while also achieving the highest balanced accuracy among the evaluated specifications.

The combined random forest improved mean accuracy by 7.2 percentage points relative to the engagement-only random forest and by 9.8 percentage points relative to the diversity-context-only random forest. Its macro AUC exceeded the corresponding reduced models by .053 and .077, respectively. As illustrated in Figure 2, the combined feature set produced consistently higher accuracy across all three algorithms.

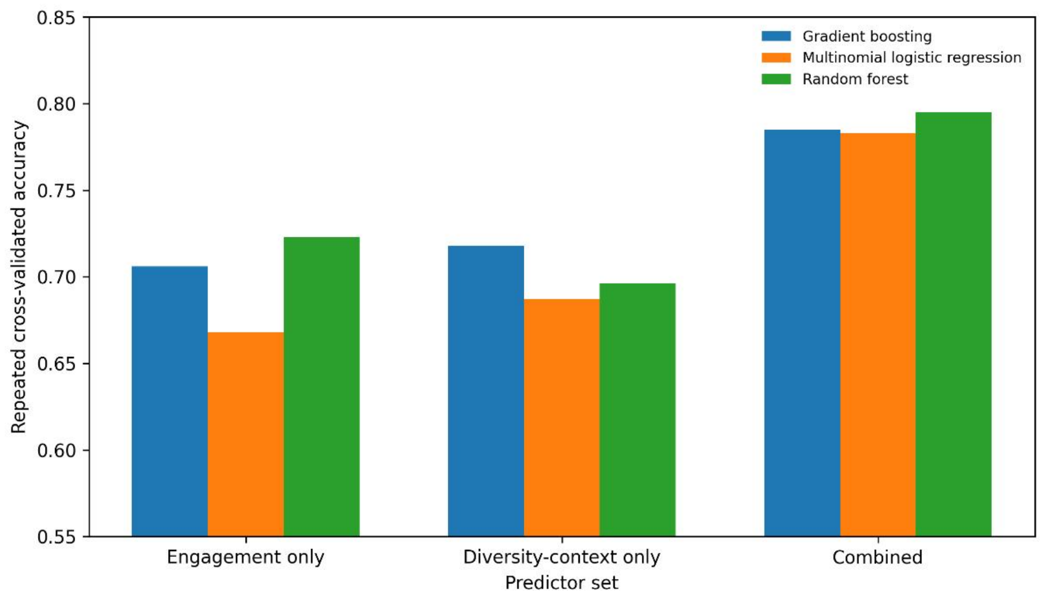

Figure 2. Cross-Validated Accuracy by Feature Set and Predictive Algorithm

Figure 2 shows that the inclusion of both engagement and learner-context variables increased predictive accuracy for multinomial logistic regression, random forest, and gradient boosting. The highest observed mean accuracy was obtained by the combined random-forest specification.

4.5 Final Model Classification Diagnostics

The combined random forest was evaluated further using one complete set of stratified five-fold out-of-fold predictions. The resulting accuracy was .794, balanced accuracy was .788, macro F1 was .797, and one-versus-rest macro AUC was .927. Table 5 presents the out-of-fold confusion matrix.

(ISSN: 2208-2115)
Vol. 12, Iss. 02, 2026
DOI:10.65327/es.v12i2.2535

Table 5. Out-of-Fold Confusion Matrix for the Combined Random Forest

Actual category	Predicted high	Predicted low	Predicted middle	Total
High	65	0	22	87
Low	0	49	13	62
Middle	15	8	109	132

Table 5 shows that 223 of the 281 observations were classified correctly. No high-achievement observation was classified as low, and no low-achievement observation was classified as high; all errors for these two categories were assigned to the adjacent middle category.

Class-specific performance is reported in Table 6. Precision ranged from .757 for middle achievement to .860 for low achievement, while recall ranged from .747 for high achievement to .826 for middle achievement.

Table 6. Class-Specific Out-of-Fold Classification Performance

Achievement category	Precision	Recall	F1 score	Support
High	.813	.747	.778	87
Low	.860	.790	.824	62
Middle	.757	.826	.790	132
Macro average	.810	.788	.797	281
Weighted average	.797	.794	.794	281

Table 6 shows that the low-achievement category produced the highest precision and F1 score, whereas the middle-achievement category produced the highest recall. The macro and weighted averages were similar, indicating that aggregate performance was not determined exclusively by the largest achievement category.

4.6 Predictor Importance and Robustness

Permutation importance was computed from the fitted combined random forest as the mean reduction in classification accuracy following repeated random permutation of each original predictor. Table 7 presents the ten highest-ranked variables.

Table 7. Permutation Importance for the Combined Random-Forest Model

Rank	Predictor	Mean accuracy reduction	SD
1	Student absence days	.201	.020
2	Parent relationship	.117	.014
3	Resources visited	.047	.009

(ISSN: 2208-2115)
Vol. 12, Iss. 02, 2026
DOI:10.65327/es.v12i2.2535

4	Gender	.033	.005
5	Raised-hand participation	.022	.007
6	Discussion activity	.013	.005
7	Subject topic	.011	.004
8	Parental survey participation	.010	.005
9	Nationality	.008	.003
10	Announcement viewing	.006	.002

Table 7 shows that student absence days generated the largest mean decline in classification accuracy when permuted, followed by parent relationship and resource visits. Three of the four direct engagement indicators appeared among the ten highest-ranked predictors.

The superiority of the combined feature set was retained across all three algorithms and across accuracy, balanced accuracy, macro F1, and macro AUC. The combined multinomial model achieved .783 accuracy and .903 macro AUC, while the combined gradient-boosting model achieved .785 accuracy and .911 macro AUC; both remained below the combined random forest on accuracy and macro AUC. Model-performance variability across repeated validation folds was limited, with standard deviations of .033 for random-forest accuracy and .017 for macro AUC.

Sensitivity analysis using separate binary achievement thresholds retained a significant positive engagement coefficient for both middle-or-high versus low achievement and high versus low-or-middle achievement. The estimated engagement odds ratios were 32.02 and 4.55, respectively, with both $p<.001$. Fewer than seven absence days also remained significant at both thresholds, with odds ratios of 26.27 and 16.66, respectively. These threshold-specific estimates preserved the direction of the corresponding ordinal-model coefficients.

5. Discussion

The results indicate that digital learning engagement is related to other components of attendance, home engagement, student characteristics, and instructional involvement and engagement that are tied to school success. The monotonic trend in the engagement index for each low, middle, and high achievement group suggests that students who engaged with the resources, announcements, discussions, and participatory activities more often were more likely to be in higher achievement groups. This is similar to the findings of digitally mediated environments as a means of enhancing participation where engagement is intentionally nurtured through deliberate learning processes (Aldhafeeri & Alotaibi, 2022). The results in the present study, however, extend this evidence by demonstrating that engagement remained a useful predictor even when this was done to account for the context of the individual's experience, and helped to classify new cases.

Theoretically important is the superiority of the combined model over engagement-only and context-only specifications. It implies that achievement is not simply related to mere use of a platform but to a combination of behavioral and contextual factors that interact to produce achievement. This reinforces the idea that a good quality, embedded technology is more important than access alone or frequency of use (Consoli et al., 2025). The variables of resource visits and raised hand participation were significant discriminative factors and viewed as more informative than announcement viewing, suggesting that active, task-oriented behaviors are more discriminative than passive behaviors. The significant impact of absence variables and parental relationship variables also points to a broader continuum of continuity of learning and support that take place when engaging with digital.

Results from both the ordinal regression and machine-learning models elucidate the difference between explanation and prediction. The regression model used engagement, gender, parental involvement and attendance as factors with significant contributions to a higher achievement while the random forest model used nonlinearities and interactions that provided better classification. This two-pronged approach is driven by evidence that engagement-achievement relationships exist across different student profiles and achievement states, and are not identical across all populations (Saqr et al., 2023). The results of the study thus provide evidence in favour of an explanatory-predictive approach based on interpretable coefficients that provide the direction of association, and algorithms that are validated across different datasets and assess classification performance.

In the context of educational technology, the results of the combined random forest showed good performance, suggesting that LMS-related data can be used to complement the data obtained from the assessment. Repeated stratified validation of reduced and combined feature sets has been performed to add value to the similar predictive effort that showed that patterns of LMS interaction can be estimated analytically (Alshammari, 2024). There are specific interest and relevance of the high precision and recall in the classification of low achievement scores to the early support, however, the model should not be used as a stand alone decision system. It should be used as a screening tool to remind teachers, remind students to attend, remind to seek resources, and to interact with families.

The study is methodologically valuable, as it is a one-stop analysis of methodologically combining construct validation, ordinal modelling, comparative machine learning, class-sensitive metrics and permutation importance applied to a secondary-education dataset. This helps alleviate concerns about one's overall accuracy and the possibility of poor performance on minority classes, or poor interpretability (Lam et al., 2024). However, there are a number of limitations to making inferences. The sample was limited and comprised mainly of middle school observations, and the data were limited to one dataset. Achievement was captured in a categorical manner instead of in terms of a score, and the data did not include socioeconomic status, disability, language proficiency, digital access, or instructional-design variables. The cross sectional structure prevented temporal ordering and causal inference, and it was not possible to determine if there was any dependence between observations since there was no student identifier available.

Future research should replicate and expand on results to make the model applicable to larger secondary-school systems, include longitudinal engagement trajectories, and determine if the predictive performance is stable over time. External studies should contain precise assessment scores, more detailed inclusion variables and audits for fairness with regard to demographic grouping. Where larger samples are warranted to enable more complex deep-learning methods, it can be explored, but with a clear comparison to interpretable methods (Alnasyan et al., 2024). Future studies should explore the impact of these models, when used under the guidance of a teacher and with ethical guidance, on support decisions and outcomes in education.

6. Conclusion

This study demonstrated the combined effect of digital learning engagement, learners' characteristics, attendance, and parents' context is more predictive of achievement than any of the predictor domains alone. The engagement construct was validated as consistently related to higher levels of achievement, and the comparative modelling process indicated that use of both behaviour and context enhanced classification accuracy and yielded the most reliable identification of academically vulnerable students. The results contribute to the educational technology field and its research by confirming that activity in LMS is not to be read as an educational process per se but as part of a larger educational ecology. In practice, schools should embed engagement measures into their own systems of monitoring participation and attendance, and also to systems of review which are led by the teacher and supported by modelling based on the output of engagement measures; they should give priority to the provision of interventions for access to resources, participation, continuity of attendance and communication with parents. To minimize the potential for misclassification and inequitable use, institutional implementation should include clear decision rules, privacy protection, subgroup performance monitoring and professional oversight. The study is limited by its categorical definition of achievement, its small sample size, and underrepresentation of high schools, as well as by its cross-sectional nature and limited socioeconomic, disability, language and digital-access factors. The framework should be replicated with larger and more diverse secondary-school populations, the engagement experiences should be longitudinal, and the framework should be refined to include finer-grained achievement outcomes and predictive fairness should be evaluated across learner groups. Future research should also examine the effects of such models on targeted support, engagement and academic growth during extended instructional periods when used in an ethical, teacher-mediated fashion.

References

1. Aldhafeeri, F. M., & Alotaibi, A. A. (2022). Effectiveness of digital education shifting model on high school students' engagement. Education and Information Technologies, 27(5), 6869-6891.

2. Alnasyan, B., Basheri, M., & Alassafi, M. (2024). The power of Deep Learning techniques for predicting student performance in Virtual Learning Environments: A systematic literature review. Computers and Education: Artificial Intelligence, 6, 100231.

3. Alshammari, A. (2024). Using analytics to predict students' interactions with learning management systems in online courses. Education and Information Technologies, 29(15), 20587-20612.

4. Ameloot, E., Rotsaert, T., Ameloot, T., Rienties, B., & Schellens, T. (2024). Supporting students' basic psychological needs and satisfaction in a blended learning environment through learning analytics. Computers & Education, 209, 104949.

5. Amrieh, E. A., Hamtini, T., & Aljarah, I. (2016). Students' academic performance dataset (xAPI-Edu-Data) [Data set]. Kaggle. https://www.kaggle.com/datasets/aljarah/xAPI-Edu-Data

6. Bakır-Yalçın, E., & Usluel, Y. K. (2024). Investigating the antecedents of engagement in online learning: do achievement emotions matter?. Education and Information Technologies, 29(4), 3759-3791.

7. Bergdahl, N. (2022). Engagement and disengagement in online learning. Computers & Education, 188, 104561.

8. Cabı, E. (2025). The relationship between academic performance and the interaction patterns of online students in learning management system. Education and Information Technologies, 30(7), 8473-8493.

9. Cigdem, H., Ozturk, M., Karabacak, Y., Atik, N., Gürkan, S., & Aldemir, M. H. (2024). Unlocking student engagement and achievement: The impact of leaderboard gamification in online formative assessment for engineering education. Education and Information Technologies, 29(18), 24835-24860.

10. Consoli, T., Schmitz, M. L., Antonietti, C., Gonon, P., Cattaneo, A., & Petko, D. (2025). Quality of technology integration matters: Positive associations with students' behavioral engagement and digital competencies for learning. Education and Information Technologies, 30(6), 7719-7752.

11. Emerson, A., Min, W., Azevedo, R., & Lester, J. (2023). Early prediction of student knowledge in game-based learning with distributed representations of assessment questions. British Journal of Educational Technology, 54(1), 40-57.

12. Fidan, A., & Koçak Usluel, Y. (2024). Emotions, metacognition and online learning readiness are powerful predictors of online student engagement: A moderated mediation analysis. Education and Information Technologies, 29(1), 459-481.

13. Lam, P. X., Mai, P. Q. H., Nguyen, Q. H., Pham, T., Nguyen, T. H. H., & Nguyen, T. H. (2024). Enhancing educational evaluation through predictive student assessment modeling. Computers and Education: Artificial Intelligence, 6, 100244.

14. Maričić, M., & Lavicza, Z. (2024). Enhancing student engagement through emerging technology integration in STEAM learning environments. Education and Information Technologies, 29(17), 23361-23389.

15. Mutimukwe, C., Viberg, O., Oberg, L. M., & Cerratto-Pargman, T. (2022). Students' privacy concerns in learning analytics: Model development. British Journal of Educational Technology, 53(4), 932-951.

16. Nazaretsky, T., Ariely, M., Cukurova, M., & Alexandron, G. (2022). Teachers' trust in AI-powered educational technology and a professional development program to improve it. British journal of educational technology, 53(4), 914-931.

17. Rodríguez-Martínez, J. A., González-Calero, J. A., del Olmo-Muñoz, J., Arnau, D., & Tirado-Olivares, S. (2023). Building personalised homework from a learning analytics based formative assessment: Effect on fifth-grade students' understanding of fractions. British Journal of Educational Technology, 54(1), 76-97.

18. Saqr, M., López-Pernas, S., Helske, S., & Hrastinski, S. (2023). The longitudinal association between engagement and achievement varies by time, students' profiles, and achievement state: A full program study. Computers & Education, 199, 104787.

19. Sarker, S., Paul, M. K., Thasin, S. T. H., & Hasan, M. A. M. (2024). Analyzing students' academic performance using educational data mining. Computers and Education: Artificial Intelligence, 7, 100263.

20. Wang, Y. J., Gao, C. L., & Ye, X. D. (2024). A data-driven precision teaching intervention mechanism to improve secondary school students' learning effectiveness. Education and Information Technologies, 29(9), 11645-11673.

21. Zheng, L., Niu, J., & Zhong, L. (2022). Effects of a learning analytics-based real-time feedback approach on knowledge elaboration, knowledge convergence, interactive relationships and group performance in CSCL. British Journal of Educational Technology, 53(1), 130-149.