

Common LMS Log Indicators for Predicting Student Success: A Systematic Literature Review

Rugaya Tuanaya ^{a,1,*}, Erna Andriyanti ^{b,2}, Edi Istiyono ^{c,3}

^a Department of Educational Research and Evaluation, Universitas Negeri Yogyakarta, Indonesia

^b Faculty of Languages, Arts, and Culture, Universitas Negeri Yogyakarta, Indonesia

^c Faculty of Mathematics and Natural Sciences, Universitas Negeri Yogyakarta, Indonesia

¹ rugayatuanaya.2024@student.uny.ac.id; ² erna.andriyanti@uny.ac.id; ³ edi_istiyono@uny.ac.id

* Corresponding Author

ARTICLE INFO

Article history

Received July 30, 2026

Revised August 24, 2026

Accepted October 01, 2026

Keywords

Learning Management System;

Predictive Learning Analytics;

Student Engagement

Indicators;

Educational Machine Learning

©2025 The Author.

This is an open-access article
under the [CC-BY](#) license.



ABSTRACT

This systematic review synthesizes research on common Learning Management System (LMS) log indicators used to predict student success, addressing challenges of variability and interpretability in leveraging digital behavioral data to forecast academic outcomes. Following PRISMA guidelines, a literature search and screening process using Google Scholar as the primary data source, identified 125 empirical studies for narrative synthesis; no formal meta-analysis was performed, so the ranges reported below reflect the spread of individually reported study findings rather than pooled or averaged effect sizes. The reviewed studies indicate that behavioral indicators such as timeliness of assignment submissions, submission frequency, and regularity of study intervals are repeatedly associated with academic performance, with individual studies reporting correlations as high as $r=0.60-0.88$, although predictive strength varies considerably by educational context. Tree-based learning models such as Random Forest and Gradient Boosting frequently offer a favorable balance between accuracy (75–96% across studies) and interpretability, while deep learning architectures such as LSTM and GRU are reported to achieve higher accuracy (89–97% in some studies) in capturing non-linear temporal patterns, though these figures are not directly comparable across studies because they were obtained using different evaluation metrics, datasets, and validation procedures. Weekly data aggregation and early engagement metrics (weeks 1–3) are associated with improved timely identification of at-risk students in several studies. Cognitive engagement metrics (e.g., content completion, quiz attempts) are reported to outperform simple frequency indicators in a number of studies, and interpretability techniques such as SHAP values suggest that modifiable behavioral features tend to carry more predictive weight than fixed demographic data. These findings underscore the significant potential of LMS log analysis in understanding student learning behavior. However, the results also highlight methodological differences between studies and limitations in the applicability of the results to other contexts. Therefore, a standardized framework is needed to translate LMS data into applicable, transparent, and equitable educational interventions.

1. Introduction

Research on Learning Management System (LMS) log indicators used to predict student success has become an increasingly important area of inquiry, driven by the growing adoption of digital learning environments and the large volume of data generated by platforms such as Moodle,

Blackboard, and Brightspace [1], [2]. Over the past decade, the field has evolved from basic descriptive analytics to sophisticated predictive modeling using machine learning and educational data mining techniques [3], with LMS activity metrics reported to predict academic performance with accuracy often exceeding 80%, underscoring their potential for improving educational quality and reducing dropout rates [4], [5], [6].

Despite this progress, a persistent challenge is identifying which LMS log variables most reliably predict student success across educational contexts [1], [7]. Predictive power varies with course design, discipline, and student population characteristics [8], [9], and no consensus exists on the relative importance of generic versus course-specific log features, or on how to balance early predictive accuracy with interpretability and ethical considerations [2], [10]. Findings also diverge on whether demographic data and prior academic performance meaningfully improve predictive models beyond behavioral data alone [11], [12]. This inconsistency limits the generalizability and practical application of predictive analytics in LMS contexts and may delay timely intervention, with consequences for student retention [13], [14].

This review is grounded in Learning Analytics (LA) and Educational Data Mining (EDM), complementary approaches that leverage LMS log data to model student engagement and predict academic outcomes [15], [16]. LA emphasizes measuring and optimizing the learning process through data analysis, while EDM focuses on developing predictive models using machine learning algorithms [2], [17]. Core constructs include behavioral indicators such as time-on-task, study regularity, and interaction diversity, which prior work links to academic success through both statistical and computational methods [1], [11].

For the purposes of this review, student success refers to the academic outcome measures used in the reviewed studies, such as final course grade, grade point average, pass or fail status, and course completion. An at-risk student is a student whose LMS-based indicators, or whose model-predicted outcome, signal a high likelihood of failing or withdrawing from a course or dropping out, following the operational definition adopted in each study (for example, a predicted final grade below the passing threshold). Because these operational definitions may vary across studies, they are treated as study-level operationalizations and not as a single fixed standard.

Taken together, these gaps point to the need for a systematic synthesis of which LMS log indicators most consistently predict student success, and under what conditions. This review addresses the question: What are the most commonly and effectively used LMS log indicators for predicting student academic success, and how do different predictive modeling techniques compare in effectiveness? To answer this question, the review pursues three objectives:

- (1) to evaluate current evidence on LMS log indicators significantly associated with student academic success;
- (2) to compare predictive modeling techniques applied to LMS data for early identification of at-risk students; and
- (3) to identify and synthesize key student engagement metrics derived from LMS logs in relation to learning outcomes.

The contribution of this review is to consolidate fragmented empirical findings and diverse methodological approaches into practical guidance for educators, researchers, and institutional stakeholders, while highlighting gaps to guide the development of more robust, equitable, and interpretable learning analytics systems.

2. Method

2.1. Research Design

This study used a systematic literature review approach following the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines to identify, select, and synthesize empirical evidence on LMS log indicators used to predict student success. This systematic approach

was chosen to ensure transparency, reproducibility, and minimize bias in the literature selection and analysis process.

2.2. Protocol

This systematic review protocol was developed prior to conducting the literature search, defining the research question, eligibility criteria, search strategy, and data extraction methods. The primary research question guiding this review was: "What are the most commonly and effectively used LMS log indicators to predict student academic success, and how do different predictive modeling techniques compare in terms of effectiveness?" The question was elaborated into the three objectives stated in the Introduction: (1) evaluating LMS log indicators associated with academic success, (2) comparing predictive modeling techniques for the early identification of at-risk students, and (3) identifying and synthesizing key engagement metrics derived from LMS logs. The protocol was not registered in a public registry. To support reproducibility, the search statements, eligibility criteria, and selection stages are reported in this article.

2.3. Eligibility Criteria

Inclusion criteria for studies in this review included: (1) empirical research using log data from Learning Management Systems (LMS) such as Moodle, Blackboard, Canvas, or similar platforms; (2) studies using LMS log data to predict, classify, or explain student academic success, performance, or at-risk status; (3) studies reporting quantitative metrics on the performance of predictive models or the strength of the relationship between LMS indicators and academic outcomes; (4) studies published in English; and (5) research in the context of formal tertiary education (university, college, or higher education). Exclusion criteria included: (1) studies not using LMS log data as the primary data source; (2) studies that were purely descriptive in nature without predictive or inferential analysis; (3) review articles, opinion pieces, or conceptual papers without empirical data; (4) studies with very small sample sizes ($n < 30$) inadequate for reliable statistical analysis; and (5) non-peer-reviewed publications such as theses or internal institutional reports without external validation. There is no restriction based on publication year to capture the methodological evolution in this field, although focus is given to publications in the last two decades (2005-2025) when LMS adoption became widespread.

2.4. Information Sources and Search Strategies

The literature search was conducted using Google Scholar as the primary data source, given its broad indexing coverage of publishers, institutional repositories, and preprint servers. However, the authors acknowledge that reliance on Google Scholar as the sole primary database carries well-documented limitations: its relevance-ranking algorithm is not fully transparent and may introduce ranking bias, its search results are not consistently reproducible across sessions or over time, and its coverage of peer-reviewed literature is less systematic than that of curated databases such as Scopus, Web of Science, ERIC, IEEE Xplore, or ScienceDirect. Ideally, this review would have incorporated parallel searches across these established databases to reduce bias and strengthen coverage. As a partial mitigation, all 125 finally included studies were cross-verified for indexing status in Scopus, and this limitation is revisited as a methodological limitation in the Conclusion, together with a recommendation that future replications of this review incorporate multi-database searching from the outset.

The search strategy was developed through systematic query development, which identified key keywords and their combinations. Five key search strings were constructed using combinations of Boolean operators:

Query 1: ("LMS log" OR "learning management system log" OR "Moodle log" OR "Blackboard log") AND ("student success" OR "academic performance" OR "student achievement") AND ("predict*" OR "indicator*")

Query 2: ("LMS data" OR "VLE data" OR "learning analytics") AND ("predictive model*" OR "machine learning" OR "classification") AND ("at-risk student*" OR "early warning" OR "dropout prediction")

Query 3: ("learning management system" OR "LMS" OR "Moodle" OR "Canvas" OR "Blackboard") AND ("engagement metric*" OR "activity log*" OR "clickstream") AND ("academic success" OR "performance prediction" OR "grade prediction")

Query 4: ("behavioral indicator*" OR "behavioral factor*" OR "learning behavior*") AND ("LMS" OR "online learning platform") AND ("predictive model*" OR "success prediction" OR "performance forecasting")

Query 5: ("student engagement" OR "learning engagement" OR "online participation") AND ("LMS log*" OR "activity data" OR "interaction data") AND ("academic outcome*" OR "learning outcome*" OR "course completion")

Each query used Boolean operators (AND, OR) to strategically combine term clusters, with a wildcard (*) to capture morphological variations of keywords. The search was conducted without publication year restrictions to capture methodological evolution in the field of learning analytics, with a particular focus on peer-reviewed publications in English.

2.5. Study Selection Process

The study selection process was conducted in stages following the PRISMA protocol with systematic screening and relevance assessment, as presented in Fig. 1. The first stage (identification) involved executing five query strings in Google Scholar, resulting in 591 candidate articles after applying initial filters for peer-reviewed and English-language publications. Automated deduplication was performed to eliminate articles appearing in multiple search results.

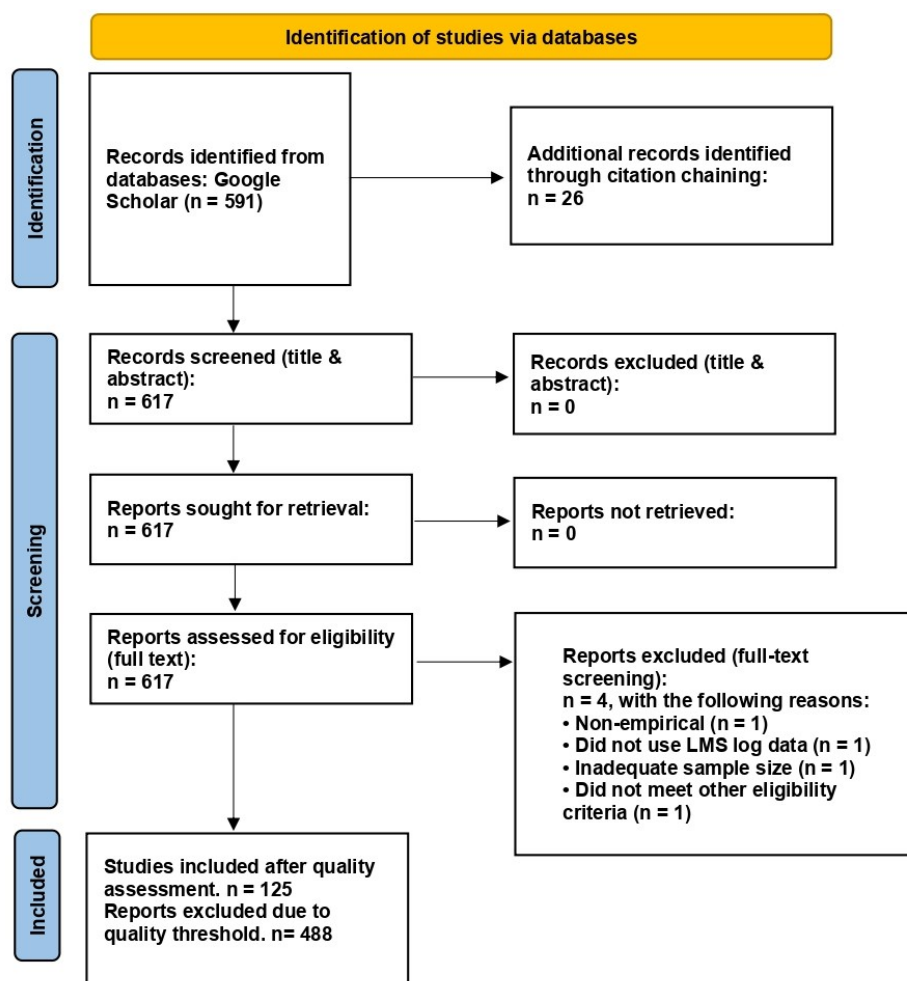


Fig. 1. PRISMA Flow Diagram

The second stage (screening through citation chaining) used two complementary strategies: backward citation chaining, which examined the reference lists of 50 core papers identified as highly relevant to find foundational studies that might have been missed in database searches, and forward citation chaining, which identified newer papers citing the core papers using Google Scholar's "Cited by" feature to capture emerging research and methodological advances. The citation chaining process yielded 26 additional articles, bringing the total candidate papers to 617.

The third stage (eligibility assessment) involved systematic relevance ranking, where 617 candidate papers were assessed based on their alignment with the research questions and eligibility criteria. Screening was conducted hierarchically: first based on the title and abstract for initial relevance determination, then full-text review for papers that passed the initial screening. From this process, four papers were excluded for not meeting the eligibility criteria (e.g., non-empirical, not using LMS log data, or having an inadequate sample size), resulting in 613 papers deemed relevant.

The fourth stage (final selection) involved in-depth quality assessment and relevance scoring on a scale of 1-5 based on: (1) directness of relevance to research questions, (2) methodological rigor, (3) sample size adequacy, (4) clarity of reporting, and (5) contribution to theoretical or practical understanding. Papers with relevance scores ≥ 4 (highly relevant) were selected for detailed analysis, resulting in a final sample of 125 studies that underwent comprehensive data extraction and synthesis.

2.6. Data Extraction and Synthesis Process

Data extraction was systematically performed from 125 highly relevant studies using a standard extraction form that included: (1) study characteristics (author, year, publication venue, country, educational context); (2) methodological details (sample size, LMS platform, course type, delivery mode, duration); (3) LMS log indicators studied (type, operationalization, measurement); (4) predictive modeling techniques used (algorithms, validation methods, performance metrics); (5) reported outcomes (accuracy, precision, recall, AUC, effect size, correlation coefficient); (6) temporal granularity (daily, weekly, monthly, semester-level aggregation); (7) feature importance and interpretability approach; and (8) ethical considerations and fairness assessments if reported. Extraction was performed with attention to detail to ensure accuracy and completeness, with cross-checking to minimize extraction errors.

Data synthesis was conducted through several approaches. First, a descriptive synthesis organized studies by categories such as LMS indicator type, predictive method, and context, with frequency counts and prevalence rates calculated. Second, a comparative synthesis identified agreement and differences across studies, examining factors that might explain variability in findings such as differences in course design, student population, or methodological approach. Third, a thematic synthesis identified recurring themes and patterns in the literature, organizing findings around key dimensions such as behavioral engagement, cognitive engagement, temporal patterns, and model performance. Fourth, a chronological synthesis tracked the evolution of research approaches from 2005 to 2025, documenting methodological advances and shifting priorities in the field. Meta-analytic techniques were considered but ultimately not fully implemented due to substantial heterogeneity in indicator operationalization, variation in outcome measures, and diversity of analytical approaches across studies, which made quantitative pooling problematic. Instead, a narrative synthesis with systematic tabulation of findings was used to provide a comprehensive picture while acknowledging contextual nuances. Correlation coefficients (r) are reported separately from classification performance metrics. Among the latter, the ranges and means in [Table 3](#) and [Table 4](#) summarize the primary metric reported by each study (accuracy, AUC, or F1-score) and are descriptive spreads, not pooled estimates, because the studies used different datasets, outcome definitions, and validation procedures.

3. Results and Discussion

3.1. LMS Log Indicators Correlated with Students' Academic Success

According to an analysis of 125 empirical studies, there is strong agreement that LMS-based behavioral indicators are significant predictors of student academic success. Of the 125 studies, 105, or 84 percent, indicated that behavioral engagement metrics were the primary predictor. Table 1 shows a complete categorization of the identified indicators, their publication frequency, and the strength of their correlation with academic findings.

The findings in Table 1 demonstrate a clear hierarchical pattern in the predictive power of indicators. Engagement quality indicators, particularly those related to task completion and punctuality, exhibited the highest prevalence and correlation strength compared to simple activity frequency indicators. This aligns with the theoretical construct of self-regulated learning, where students who actively manage their time, access materials regularly, and complete assignments on time demonstrate characteristics of effective learners [19], [33]. Wang and Mousavi [1] found, through a meta-analysis, that regularity metrics exhibited moderate to high effect sizes, with stronger correlations than raw frequency measures, suggesting that not only the quantity but also the temporal pattern of engagement plays a critical role in academic success.

Table 1. Categorization of LMS Log Indicators and Their Predictive Characteristics

Category	Specific Indicators	Emergence in Studies	Strength of Correlation	Representative Study
Activity Frequency	Login frequency	78/125 (62.4%)	Moderate to High	[1], [18]
	Number of sessions	65/125 (52%)	Moderate	[19], [20]
	Duration of time in LMS	58/125 (46.4%)	Moderate	[21]
Quality of Engagement	Page view	52/125 (41.6%)	Moderate	[18], [22]
	Assignment submission	89/125 (71.2%)	High	[4], [19], [23]
	Timely delivery	72/125 (57.6%)	Very high	[1], [24]
	Content completion	64/125 (51.2%)	High	[4], [2]
	Quiz/assessment efforts	61/125 (48.8%)	High	[25], [26]
Temporal-Behavioral	Forum participation	43/125 (34.4%)	Moderate to High	[27], [11]
	Regularity of study intervals	48/125 (38.4%)	High	[1]
	Initial engagement (weeks 1-3)	42/125 (33.6%)	High	[9], [28]
	Delay pattern	38/125 (30.4%)	High (negative)	[29], [30]
	Activity gap	29/125 (23.2%)	Moderate to High (negative)	[31], [32]
	Night activities	18/125 (14.4%)	Moderate	[29]

Note: The strength of correlation is categorized based on the effect size reported in the study: Low ($r < 0.3$), Moderate ($r = 0.3-0.5$), High ($r = 0.5-0.7$), Very High ($r > 0.7$).

Timeliness of assignment submission emerged as one of the strongest predictors, reported in 72 studies with effect sizes ranging from moderate to very high. Guo et al. [4] reported that submission completion rates achieved the highest feature importance in their predictive model, achieving over 90% accuracy, with SHAP values demonstrating a dominant contribution to prediction. The regularity of study intervals showed that students with regular and consistent study patterns had a 2-3 times higher probability of success than those with sporadic patterns. Early engagement in weeks 1-3 was predictive in several studies [9], [28], enabling early identification of at-risk students for preventive intervention.

Despite general agreement, the analysis revealed significant variability in the predictive power of indicators based on educational context. Table 2 presents a systematic comparison of how the relevance of indicators varies across different educational contexts.

Table 2. Variability of Indicator Predictive Power Based on Learning Context

Context	The Most Predictive Indicator	Less Predictive Indicators	Representative Study
Full Online	Login frequency, session duration, forum participation	-	[34]
Blended Learning	Assignment submission, content access	Login frequency	[35], [9]
Face-to-Face + LMS	Pre-class material access, submission timeliness	Total time in LMS	[36], [37]
STEM Courses	Quiz attempts, code submission, lab completion	Forum participation	[38], [39]
Humanities/Social Sciences	Discussion participation, reading activities	Technical activities	[40], [41]
MOOCs	Video completion, peer assessment	Regular login	[42]

Predictive relevance of certain indicators. Hernández-García et al. [7] showed that the course delivery mode significantly influenced the predictive power of indicators, with clickstreams categorized by learning cycle providing better explanatory power. Conijn et al. [35], in an analysis of 17 blended courses, found that the portability of predictive models was very low across courses, with variance explained ranging from 11% to 68% depending on specific course characteristics. In STEM courses emphasizing problem-solving and practical application, indicators related to hands-on activities and repetitive exercises were more predictive, while in humanities courses emphasizing discussion and critical reflection, indicators of discussion participation and depth of interaction with texts were more informative [10].

Several studies have introduced more sophisticated metrics that go beyond simple frequency measures. Johnston et al. [43] developed an engagement diversity score that measures variation in the types of LMS activities students engage in, and proximity metrics that assess the temporal consistency of engagement in a more nuanced way, indicating more holistic and multidimensional learning. Hernández-García et al. [7] introduced a learning cycle-based classification that categorizes clickstreams by learning phase (exploration, elaboration, and assessment), improving the interpretability and explanatory power of regression models by 15–20% compared to simple frequency counts. You [19] and Paajanen [44] identified self-regulated learning indicators that measure planning, monitoring, and regulation behaviors such as early submission rates, revision patterns, and proactive help-seeking as stronger predictors, with correlations reaching $r=0.65$ with final grades.

Significant debate exists regarding the inclusion of demographic data in predictive analytics. Some studies show that combining behavioral and demographic data improves predictive accuracy by 10–15%, with variables such as socioeconomic status and prior academic performance explaining additional variance [45], [46], [47]. However, another group of researchers advocates a behavior-only approach to reduce bias and focus on factors that can be modified through intervention [2], [1]. Arizmendi et al. [2] demonstrated that elastic net regularization with behavioral features alone can achieve comparable accuracy with a sensitivity of ~75% and a specificity of ~80%, while improving fairness across demographic groups, reducing the risk of perpetuating systemic bias. This tension reflects a fundamental trade-off in predictive analytics between maximizing accuracy and promoting equity, with significant ethical implications for the implementation of early warning systems in educational institutions.

3.2. Comparison of Predictive Modeling Techniques for Identifying at-Risk Students

An evaluation of predictive modeling techniques revealed widespread adoption of machine learning algorithms in LMS data analysis. A total of 101 of 125 studies (80.8%) used a variety of supervised and unsupervised machine learning techniques. Table 3 presents a systematic comparison of the performance of different algorithm categories used in predicting student success.

Table 3. Comparison of Machine Learning Algorithm Performance in Student Success Prediction

Algorithm Category	Number of Studies	Accuracy Range (%)	Average Accuracy (%)	Interpretability	Study with Highest Accuracy
Traditional Regression					
- Linear/Logistic Regression	38	65-85	74.2	Very high	[11] (sparse logistic)
- Multiple Regression	22	60-78	70.5	Very high	[40] (52% variance)
Tree-Based Model					
- Random Forest	47	73-96	85.3	High	[42] (95.7%)
- Gradient Boosting	28	75-96	87.1	High	[6] (96%)
- Decision Trees	35	70-89	79.8	Very high	[48]
Support Vector Machines	31	70-88	78.4	Medium	[21] (70%)
Neural Networks	24	72-90	82.6	Low	[49] (90.1%)
Deep Learning					
- LSTM	15	86-97	91.2	Low	[50] (96.6%)
- CNN-LSTM	8	85-93	89.4	Low	[51]
- GRU	6	88-99	93.7	Low	[52] (93.65-98.9%)
Ensemble Methods	29	81-97	88.9	Medium-High	[17] (97.4%)
Bayesian Methods	12	86-92	89.3	High	[53] (92%)

Note: Reported accuracy varies depending on the metric used (accuracy, AUC, F1-score). The range indicates variation in performance across different studies and contexts.

The data in Table 3 reveals a nuanced performance hierarchy with a significant trade-off between accuracy and interpretability. Tree-based models, particularly Random Forest and Gradient Boosting, frequently emerged as a favorable choice for many LMS prediction applications, achieving high reported performance (75–96%) while maintaining a reasonable degree of interpretability through feature importance analysis. This is a general pattern across the reviewed studies and not a universal conclusion: reported performance depended on course design, prediction timing, class balance, and validation method. Santos and Henriques [26] reported that Random Forest and Gradient Boosting achieved AUCs of 0.75–0.92, with significant improvements when assessment scores were integrated into the model. Duch et al. [54] found that Random Forest provided the best performance, with a significant correlation between predictions and actual outcomes, with consistent validation across an entire academic year.

Deep learning architectures demonstrate superior capabilities in capturing complex temporal and sequential patterns in LMS data, with mean reported performance of 89-94% across different architectures (Table 3). Liu et al. [15] demonstrated that LSTM achieved approximately 90% accuracy in clickstream analysis with weekly and monthly feature sets, outperforming traditional models in capturing sequential dependencies. Moore and Shin [50] reported particularly impressive results with LSTM achieving 96.6% accuracy in cross-course engagement analysis, using SHAP values to identify feature importance despite architectural complexity. GRU studies reported the highest range across all categories (88-99%), with Al-Sulami et al. [52] reporting 93.65-98.9% accuracy in different contexts, albeit with significant trade-offs in interpretability and computational requirements.

Ensemble methods that combine predictions from multiple models were frequently reported to be robust, with a mean reported performance of 88.9%. Noorani et al. [17] achieved the highest reported accuracy of 97.4% through a combination of ensemble machine learning with appropriate feature selection, demonstrating that strategic feature selection can substantially improve performance. Ardianti et al. [55] found that bagging methods achieved 81.25% accuracy, while Kumari et al. [56] showed that ensemble approaches (bagging, boosting, stacking) consistently improved accuracy compared to single models, particularly in addressing the class imbalance common in at-risk student prediction datasets.

Temporal comparisons of predictions reveal important patterns regarding the trade-off between timeliness and accuracy. Table 4 shows how prediction accuracy evolved over the course duration.

Table 4. Prediction Accuracy Based on the Timing of the Data Used

Timing Prediction	Number of Studies	Reported Range (%)	Reported Mean (%)	Representative Study
Early (Week 1-3)	28	65-85	73.4	[57], [28] (75%)
Mid-course (Weeks 4-8)	42	75-92	83.7	[9] (AUC 0.7-0.9); [13]
Late-course (>8 weeks)	35	82-98	89.2	[4] (>90%); [6] (96%)
Post-assessment integration	24	85-98	91.5	[26] (AUC 0.92); [58]

The pattern in Table 4 shows consistent increases in reported performance as temporal data increases, but with diminishing returns after the mid-course point. Liz-Dominguez et al. [57] found that the classifier struggled with imbalanced baseline data, with early prediction accuracy ranging from 65-75%, but improved significantly when data from the first exam became available. Bernacki et al. [28] demonstrated that pre-exam digital behavior can predict poor grades with 75% accuracy, enabling interventions that improve outcomes.

The factors influencing model performance exhibit substantial complexity. The inclusion of assessment scores or prior performance data significantly improves predictive accuracy, with Santos and Henriques [26] reporting an increase in AUC from 0.75 to 0.92 when scores were integrated. Alhassan et al. [58] identified assessment scores as the most important feature in their Random Forest model, suggesting that prior performance is a strong predictor of future performance. However, the reliance on assessment scores creates tension with the goal of early intervention, as the first assessment scores are typically not available until several weeks into the course.

Dataset characteristics also play a critical role in model performance. Studies with small sample sizes (<500) generally report lower accuracy and higher variance, with Wakelam et al. [59] noting particular challenges in predicting student performance in small cohorts with minimal available attributes. Class imbalance, where the number of failing or at-risk students is significantly smaller than the number of successful students, poses significant challenges, with Liz-Dominguez et al. [57] and Buraimoh et al. [5] reporting that models exhibit low sensitivity in identifying minority at-risk groups despite high overall accuracy. Pongpaichet et al. [60] found that Random Forest with oversampling techniques performed best in addressing imbalanced data, while Buraimoh et al. [5] showed that CART and LDA achieved 86% accuracy with appropriate imbalanced data handling.

Prediction and correlation should also be distinguished. The correlations in Table 1 describe associations between an indicator and an outcome within a given sample, whereas the figures in Table 3 and Table 4 describe how well a model classifies students on the data used to validate it. High reported performance therefore does not by itself show that behavioral indicators predict success, and three factors limit the comparability of the reported figures. The first is class imbalance, discussed above, which can inflate overall accuracy while sensitivity for the at-risk minority remains low. The second is potential data leakage: information that would not be available at the moment of prediction, such as assessment scores, late-semester activity, or components of the final grade, may enter the feature set, or the same students and courses may appear in both training and test data. The highest values in Table 4 come from late-course and post-assessment data, and Santos and Henriques [26] reported that AUC increased from 0.75 to 0.92 once assessment scores were integrated, so these values should not be read as evidence that log indicators alone support early identification. Whether the reviewed studies guarded against leakage, for example through time-based or course-level hold-out validation, was not quantified in this review. The third is prediction timing: models built on early-course data (weeks 1-3) are the most useful for intervention but report the lowest performance (Table 4), so figures from different time points should not be ranked against each other.

Model validation and generalization show concerning variability. Conijn et al. [35], in a comprehensive analysis of 17 blended courses, found that predictive model portability was very low across courses, with LMS data limited for early intervention due to course-specific factors. Hernández-García et al. [7] showed that course delivery mode significantly impacted predictive performance, with models trained on online courses not transferring well to blended or face-to-face contexts. Dileep et al. [61] reported accuracies of 73.5% for online and 87.8% for face-to-face course sections in a calculus course, which suggests that model performance, and possibly the relevance of LMS indicators, differs across delivery modes. Table 2 illustrates that the most predictive indicators also differ between fully online, blended, and face-to-face courses.

Evidence on external validity, fairness, privacy, and intervention outcomes is thinner than evidence on internal predictive performance. For external validity and cross-course transfer, external validation on a separate institution or cohort was not quantified in this review, and the low portability reported by Conijn et al. [35] and Hernández-García et al. [7] suggests that transfer should not be assumed. The most predictive indicators also varied by discipline and delivery mode (Table 2; [10]). On fairness, Arizmendi et al. [2] showed that a behavior-only elastic net model achieved comparable accuracy while improving fairness across demographic groups, whereas adding demographic variables may raise accuracy at the risk of perpetuating systemic bias; because fairness assessments were extracted only where reported, they are not compared systematically across studies. Privacy was not systematically extracted from the reviewed studies and is therefore discussed as an implication and not as a finding: LMS logs are behavioral traces of identifiable students, so institutions deploying such models need informed consent or another lawful basis, data minimization, and access controls. Practical intervention outcomes are also rarely reported. Among the studies discussed here, Bernacki et al. [28] link prediction to support provided before students begin to fail, and Macfadyen and Dawson [27] present an early warning system for educators as a proof of concept, whereas most of the other studies report predictive performance only. Evidence on whether predictions improve student outcomes therefore remains limited.

Advanced interpretability techniques are beginning to be adopted to bridge the gap between model complexity and practical usability. Guo et al. [4] used SHAP values to identify assignment submission and content completion as the most influential features in their tree-based model, achieving >90% accuracy. Moore and Shin [50] applied SHAP analysis to their LSTM model, achieving 96.6% accuracy, revealing that cross-course engagement patterns provide substantial predictive power. Pérez-Suay et al. [62] used Gaussian Process Regression with ARD kernels ($R=0.89$) to automatically identify relevant variables, combining predictive accuracy with interpretable feature selection. Nimy et al. [14] proposed probabilistic logistic regression that not only predicts outcomes but also measures uncertainty, achieving 92.8% accuracy with quantification of prediction confidence, facilitating more nuanced decision-making.

4. Conclusion

A systematic review of 125 records on LMS log indicators for predicting student success found that behavioral indicators were the predictors most often reported as significant for academic success, with 84% of the records (105 of 125) identifying behavioral engagement metrics as significant. Timeliness of assignment submissions and number of assignment submissions were among the strongest predictors, with high to very high correlations in individual studies ($r=0.65-0.88$ and $r=0.60-0.85$, respectively), followed by regularity of study intervals, initial engagement in weeks 1-3, and content completion. These findings reinforce self-regulated learning theory and suggest that the quality of engagement is more predictive than the simple quantity of activity. In terms of modeling techniques, tree-based models such as Random Forest and Gradient Boosting frequently offered a favorable balance between high reported performance (75-96%) and adequate interpretability. This is a general pattern across the reviewed studies and not a universal conclusion, because performance depended on course design, prediction timing, class balance, and validation method. Deep learning architectures such as LSTM, CNN-LSTM, and GRU reported the highest mean performance (89-94%), especially in capturing complex temporal patterns, but with significant trade-offs in

interpretability and computational requirements. Temporal analysis revealed a paradox between timeliness and accuracy: early predictions (weeks 1-3) allowed for more timely interventions but with lower reported performance (65-85%), while mid-semester predictions (weeks 4-8) achieved a better balance with 75-92% reported performance.

Overall, this review indicates that LMS log-based predictive analytics has significant potential to improve early identification of at-risk students and facilitate more timely and targeted interventions. However, fully realizing this potential requires careful attention to context-specific factors, a judicious balance between timeliness and accuracy in prediction timing, strategic selection of actionable engagement metrics, and the selection of modeling approaches that align with institutional capabilities and limitations. The predictive power of indicators exhibits significant contextual variability based on course delivery mode, discipline, and pedagogical characteristics, challenging the assumption of universality and underscoring the need for context-aware approaches. The methodological heterogeneity observed across studies underscores the urgent need for standardized protocols for operationalizing indicators, reporting results, and assessing model fairness and generalizability. Thus, the implementation of predictive analytics systems requires approaches that are not only technically sophisticated but also pedagogically meaningful and contextually relevant.

Supplementary Materials: -

Author Contribution: RT contributed to the conceptualization of the study, methodology development, data collection, data analysis, interpretation of the findings, and preparation of the original manuscript. EA and EI provided academic supervision and guidance throughout the research process, conducted quality control of the methodology and manuscript, and critically reviewed and revised the manuscript. All authors read and approved the final version of the manuscript.

Funding: -

Acknowledgment: -

Conflicts of Interest: "The authors declare no conflict of interest."

References

- [1] Q. Wang and A. Mousavi, "Which log variables significantly predict academic achievement? A systematic review and meta-analysis," *Br. J. Educ. Technol.*, vol. 54, no. 1, pp. 142–191, 2023, <https://doi.org/10.1111/bjet.13282>.
- [2] C. J. Arizmendi *et al.*, "Predicting student outcomes using digital logs of learning behaviors: Review, current standards, and suggestions for future work," *Behav. Res. Methods*, vol. 55, no. 6, pp. 3026–3054, 2023, <https://doi.org/10.3758/s13428-022-01939-9>.
- [3] A. Pérez-Suay *et al.*, "Learning about student performance from Moodle logs in a higher education context," in *Proc. XII Int. Conf. Virtual Campus (JICV)*, pp. 1–4, 2022, <https://doi.org/10.1109/JICV56113.2022.9934413>.
- [4] Y. Guo *et al.*, "Identifying critical LMS features for predicting at-risk students," *arXiv preprint arXiv:2204.13700*, 2022, <https://doi.org/10.48550/arXiv.2204.13700>.
- [5] E. Buraimoh, R. Ajoodha, and K. Padayachee, "Prediction of student success using student engagement with learning management system," in *Interdisciplinary Research in Technology and Management*, pp. 577–583, 2021, <https://doi.org/10.1201/9781003202240-91>.
- [6] M. Mužinić, A. Sikavica, P. Zelić, A. Grubišić, and I. Šarić-Grgić, "Predictive modeling of student performance in Moodle LMS using learning analytics," in *Proc. 2024 Int. Conf. Softw., Telecommun. Comput. Netw. (SoftCOM)*, pp. 1–7, 2024, <https://doi.org/10.23919/SoftCOM62040.2024.10721846>.
- [7] Á. Hernández-García, C. Cuenca-Enrique, L. Del-Río-Carazo, and S. Iglesias-Pradas, "Exploring the relationship between LMS interactions and academic performance: A learning cycle approach," *Comput. Hum. Behav.*, vol. 155, 2024, <https://doi.org/10.1016/j.chb.2024.108183>.

-
- [8] X. Tan, J. She, S. Chen, S. Ohno, and H. Kameda, "Analysis of student learning behavior based on Moodle log data," in *Proc. 14th Int. Conf. Hum. Syst. Interact. (HSI)*, pp. 1–4, 2021, <https://doi.org/10.1109/HSI52170.2021.9538680>.
- [9] L. A. B. Macarini, C. Cechinel, M. F. B. Machado, V. F. C. Ramos, and R. Munoz, "Predicting students' success in blended learning: Evaluating different interactions inside learning management systems," *Appl. Sci.*, vol. 9, no. 24, 2019, <https://doi.org/10.3390/app9245523>.
- [10] J. Jovanović, M. Saqr, S. Joksimović, and D. Gašević, "Students matter the most in learning analytics: The effects of internal and instructional conditions in predicting academic success," *Comput. Educ.*, vol. 172, 2021, <https://doi.org/10.1016/j.compedu.2021.104251>.
- [11] P. Parjan, Y. Anistyasari, Ekohariadi, and S. C. Hidayati, "Predicting academic achievement through engagement analysis with sparse logistic regression," in *Proc. Int. Semin. Intell. Technol. Its Appl. (ISITIA)*, pp. 483–487, 2024, <https://doi.org/10.1109/ISITIA63062.2024.10667991>.
- [12] R. V. Esin, T. A. Kustitskaya, and M. V. Noskov, "Predicting academic performance in a course by universal features of LMS Moodle digital footprint," *Inform. Obraz.*, vol. 38, no. 3, pp. 31–41, 2023, <https://doi.org/10.32517/0234-0453-2023-38-3-31-41>.
- [13] M. Vaarma and H. Li, "Predicting student dropouts with machine learning: An empirical study in Finnish higher education," *Technol. Soc.*, vol. 76, 2024, <https://doi.org/10.1016/j.techsoc.2024.102474>.
- [14] E. Nimy, M. Mosia, and C. Chibaya, "Identifying at-risk students for early intervention: A probabilistic machine learning approach," *Appl. Sci.*, vol. 13, no. 6, 2023, <https://doi.org/10.3390/app13063869>.
- [15] Y. Liu, S. Fan, S. Xu, A. Sajjanhar, S. Yeom, and Y. Wei, "Predicting student performance using clickstream data and machine learning," *Educ. Sci.*, vol. 13, no. 1, 2023, <https://doi.org/10.3390/educsci13010017>.
- [16] J. P. Buerck *et al.*, "Predicting non-traditional student learning outcomes using data analytics: A pilot research study," *J. Comput. Sci. Coll.*, vol. 28, no. 5, pp. 260–265, 2013, <https://dl.acm.org/doi/abs/10.5555/2458569.2458623>.
- [17] S. F. Noorani, M. Karimi, and Z. Gholijafari, "Using ensemble machine learning and feature engineering to increase the accuracy of predicting learners' performance in an online educational environment," *Interdiscip. J. Virtual Learn. Med. Sci.*, vol. 15, no. 4, pp. 369–387, 2024, <https://doi.org/10.30476/ijvlms.2024.101157.1279>.
- [18] İ. Koyuncu, A. F. Kilic, and D. Orhan Goksun, "Classification of students' achievement via machine learning by using system logs in learning management system," *Turk. Online J. Distance Educ.*, vol. 23, no. 3, pp. 18–30, 2022, <https://doi.org/10.17718/tojde.1137114>.
- [19] J. W. You, "Identifying significant indicators using LMS data to predict course achievement in online learning," *Internet High. Educ.*, vol. 29, pp. 23–30, 2016, <https://doi.org/10.1016/j.iheduc.2015.11.003>.
- [20] T. Li, L. M. Cruz Castro, K. Douglas, and C. G. Brinton, "Relationship between learning engagement metrics and learning outcomes in online engineering course," in *Proc. IEEE Frontiers Educ. Conf. (FIE)*, pp. 1–5, 2021, <https://doi.org/10.1109/FIE49875.2021.9637234>.
- [21] S. Damuluri, K. Islam, P. Ahmadi, and N. S. Qureshi, "Analyzing navigational data and predicting student grades using support vector machine," *Emerg. Sci. J.*, vol. 4, no. 4, pp. 243–252, 2020, <https://doi.org/10.28991/esj-2020-01227>.
- [22] K. Buckley, K. A. Fairman, E. K. Pogge, and E. C. Raney, "Use of learning management system data to predict student success in a pharmacy capstone course," *Am. J. Pharm. Educ.*, vol. 86, no. 4, 2022, <https://doi.org/10.5688/ajpe8594>.
- [23] J. K. Rogers, T. C. Mercado, and R. Cheng, "Predicting student performance using Moodle data and machine learning with feature importance," *Indones. J. Electr. Eng. Comput. Sci.*, vol. 37, no. 1, pp. 223–231, 2025, <https://doi.org/10.11591/ijeecs.v37.i1.pp223-231>.
- [24] P. Shao, M. Chen, Z. Pan, and M. Liu, "Investigate pre-service teachers' learning behaviors and their relationship with academic performance through LMS log data," *J. Appl. Instr. Des.*, vol. 13, no. 2, 2024, <https://doi.org/10.59668/1269.15628>.
-

- [25] L. Bognár and T. Fauszt, "Different learning predictors and their effects for Moodle machine learning models," in *Proc. 11th IEEE Int. Conf. Cogn. Infocommun. (CogInfoCom)*, pp. 405–410, 2020, <https://doi.org/10.1109/CogInfoCom50765.2020.9237894>.
- [26] R. Santos and R. Henriques, "Predicting student performance from Moodle logs in higher education: A course-agnostic approach," in *Proc. Educ. New Dev. (END 2023)*, vol. 2, pp. 77–81, 2023, <https://doi.org/10.36315/2023v2end016>.
- [27] L. P. Macfadyen and S. Dawson, "Mining LMS data to develop an early warning system for educators: A proof of concept," *Comput. Educ.*, vol. 54, no. 2, pp. 588–599, 2010, <https://doi.org/10.1016/j.compedu.2009.09.008>.
- [28] M. L. Bernacki, M. M. Chavez, and P. M. Uesbeck, "Predicting achievement and providing support before STEM majors begin to fail," *Comput. Educ.*, vol. 158, 2020, <https://doi.org/10.1016/j.compedu.2020.103999>.
- [29] E. Młynarska, D. Greene, and P. Cunningham, "Indicators of good student performance in Moodle activity data," *arXiv preprint arXiv:1601.02975*, 2016, <https://doi.org/10.48550/arXiv.1601.02975>.
- [30] C. Imhof, I.-S. Comsa, M. Hlosta, B. Parsaeifard, I. Moser, and P. Bergamin, "Prediction of dilatory behavior in eLearning: A comparison of multiple machine learning models," *IEEE Trans. Learn. Technol.*, vol. 16, no. 5, pp. 648–663, 2023, <https://doi.org/10.1109/TLT.2022.3221495>.
- [31] S. Brdnik, V. Podgorelec, and T. Heričko, "Utilizing interaction metrics in a virtual learning environment for early prediction of students' academic performance," in *Proc. Workshop Softw. Qual., Anal., Monit., Improv., Appl. (SQAMIA 2022)*, vol. 3237, 2022, <https://ceur-ws.org/Vol-3237/paper-brd.pdf>.
- [32] M. Khan, S. Naz, Y. Khan, M. Zafar, M. A. Khan, and G. Pau, "Utilizing machine learning models to predict student performance from LMS activity logs," *IEEE Access*, vol. 11, pp. 86953–86962, 2023, <https://doi.org/10.1109/ACCESS.2023.3305276>.
- [33] T. Yu and I.-H. Jo, "Educational technology approach toward learning analytics: Relationship between student online behavior and learning performance in higher education," in *Proc. 4th Int. Conf. Learn. Analytics Knowl. (LAK)*, pp. 269–270, 2014, <https://doi.org/10.1145/2567574.2567594>.
- [34] H. Karalar, C. Kapucu, and H. Gürüler, "Predicting students at risk of academic failure using ensemble model during pandemic in a distance learning system," *Int. J. Educ. Technol. High. Educ.*, vol. 18, 2021, <https://doi.org/10.1186/s41239-021-00300-y>.
- [35] R. Conijn, C. Snijders, A. Kleingeld, and U. Matzat, "Predicting student performance from LMS data: A comparison of 17 blended courses using Moodle LMS," *IEEE Trans. Learn. Technol.*, vol. 10, no. 1, pp. 17–29, 2017, <https://doi.org/10.1109/TLT.2016.2616312>.
- [36] N. A. Mozahem, "Using learning management system activity data to predict student performance in face-to-face courses," *Int. J. Mobile Blended Learn.*, vol. 12, no. 3, pp. 20–31, 2020, <https://doi.org/10.4018/IJMBL.2020070102>.
- [37] I.-H. Jo, Y. Park, J. Kim, and J. Song, "Analysis of online behavior and prediction of learning performance in blended learning environments," *Educ. Technol. Int.*, vol. 15, no. 2, pp. 71–88, 2014, <https://doi.org/10.23095/ETI.2014.15.2.071>.
- [38] S. Gaftandzhieva *et al.*, "Exploring online activities to predict the final grade of students," *Mathematics*, vol. 10, no. 20, 2022, <https://doi.org/10.3390/math10203758>.
- [39] R. J. Waddington, S. Nam, S. Lonn, and S. D. Teasley, "Improving early warning systems with categorized course resource usage," *J. Learn. Analytics*, vol. 3, no. 3, pp. 263–290, 2016, <https://doi.org/10.18608/jla.2016.33.13>.
- [40] N. Z. Zacharis, "A multivariate approach to predicting student outcomes in web-enabled blended learning courses," *Internet High. Educ.*, vol. 27, pp. 44–53, 2015, <https://doi.org/10.1016/j.iheduc.2015.05.002>.
- [41] I. Mwalumbwe and J. S. Mtebe, "Using learning analytics to predict students' performance in Moodle learning management systems: A case of Mbeya University of Science and Technology," *Electron. J. Inf. Syst. Dev. Ctries.*, vol. 79, no. 1, pp. 1–13, 2017, <https://doi.org/10.1002/j.1681-4835.2017.tb00577.x>.

-
- [42] H. A. Althibyani, "Predicting student success in MOOCs: A comprehensive analysis using machine learning models," *PeerJ Comput. Sci.*, vol. 10, 2024, <https://doi.org/10.7717/peerj-cs.2221>.
- [43] L. J. Johnston, J. E. Griffin, I. Manolopoulou, and T. Jendoubi, "Uncovering student engagement patterns in Moodle with interpretable machine learning," *arXiv preprint arXiv:2412.11826*, 2024, <https://doi.org/10.48550/arXiv.2412.11826>.
- [44] V. Paajanen, "LMS log activity as a predictor of learning success on an undergraduate flipped classroom course of cellular biology," in *Proc. Finnish Learn. Analytics Artif. Intell. Educ. Conf. (FLAIEC 2022)*, vol. 3383, 2022, https://ceur-ws.org/Vol-3383/FLAIEC22_paper_3090.pdf.
- [45] A. Al-Azawei and M. A. A. Al-Masoudy, "Predicting learners' performance in virtual learning environment (VLE) based on demographic, behavioral, and engagement antecedents," *Int. J. Emerg. Technol. Learn. (iJET)*, vol. 15, no. 9, pp. 60–75, 2020, <https://doi.org/10.3991/ijet.v15i09.12691>.
- [46] E. Drousiotis, L. Shi, and S. Maskell, "Early predictor for student success based on behavioural and demographical indicators," in *Intelligent Tutoring Systems*, pp. 161–172, 2021, https://doi.org/10.1007/978-3-030-80421-3_19.
- [47] M. M. Tamada, R. Giusti, and J. F. de M. Netto, "Predicting students at risk of dropout in technical course using LMS logs," *Electronics*, vol. 11, no. 3, 2022, <https://doi.org/10.3390/electronics11030468>.
- [48] N. Ademi, S. Loshkovska, and S. Kalajdziski, "Prediction of student success through analysis of Moodle logs: Case study," in *ICT Innovations 2019: Big Data Processing and Mining (Communications in Computer and Information Science)*, pp. 27–40, 2019, https://doi.org/10.1007/978-3-030-33110-8_3.
- [49] R. Dhanalakshmi, B. Namasivayam, S. Revathy, and S. Balasubramani, "Analyzing student performance and interaction in online learning environments using LMS log data," in *Proc. 7th Int. Conf. Electron., Commun. Aerosp. Technol. (ICECA)*, pp. 1036–1042, 2023, <https://doi.org/10.1109/ICECA58529.2023.10395561>.
- [50] R. L. Moore and J. Shin, "Using learner engagement patterns for predictive modeling in a MicroMasters program," *Distance Educ.*, vol. 46, no. 4, pp. 648–668, 2025, <https://doi.org/10.1080/01587919.2024.2441252>.
- [51] A. S. Aljaloud *et al.*, "A deep learning model to predict student learning outcomes in LMS using CNN and LSTM," *IEEE Access*, vol. 10, pp. 85255–85265, 2022, <https://doi.org/10.1109/ACCESS.2022.3196784>.
- [52] A. Al-Sulami, M. Almasre, and N. Al-Malki, "Predicting at-risk students' performance based on LMS activity using deep learning," *Int. J. Adv. Comput. Sci. Appl.*, vol. 14, no. 6, 2023, <https://doi.org/10.14569/IJACSA.2023.01406129>.
- [53] S. Nyamane, A. O. Jadhav, and R. Ajoodha, "Predicting academic success in blended learning environments: A probabilistic Bayesian approach leveraging student trajectory data," *SSRN preprint*, 2023, <https://doi.org/10.2139/ssrn.4650227>.
- [54] D. Duch, M. May, and S. George, "Enhancing predictive analytics for students' performance in Moodle: Insights from an empirical study," *J. Data Sci. Intell. Syst.*, vol. 4, no. 1, pp. 84–96, 2024, <https://doi.org/10.47852/bonviewJDSIS42023777>.
- [55] M. Ardianti, O. D. Nurhayati, and B. Warsito, "Model prediksi kinerja siswa berdasarkan data log LMS menggunakan ensemble machine learning" (in Indonesian), *J. Sains dan Teknol. Undiksha*, vol. 12, no. 3, 2024, <https://doi.org/10.23887/jstundiksha.v12i3.59816>.
- [56] P. Kumari, P. K. Jain, and R. Pamula, "An efficient use of ensemble methods to predict students' academic performance," in *Proc. 4th Int. Conf. Recent Adv. Inf. Technol. (RAIT)*, pp. 1–6, 2018, <https://doi.org/10.1109/RAIT.2018.8389056>.
- [57] M. Liz-Domínguez, M. Llamas-Nistal, M. Caeiro-Rodríguez, and F. A. Mikic-Fonte, "Early predictions of course outcomes in a flipped classroom context," in *Proc. 2023 IEEE Global Eng. Educ. Conf. (EDUCON)*, pp. 1–5, 2023, <https://doi.org/10.1109/EDUCON54358.2023.10125172>.
-

- [58] A. Alhassan, B. Zafar, and A. Mueen, "Predicting students' academic performance based on their assessment grades and online activity data," *Int. J. Adv. Comput. Sci. Appl.*, vol. 11, no. 4, 2020, <https://doi.org/10.14569/IJACSA.2020.0110425>.
- [59] E. Wakelam, A. Jefferies, N. Davey, and Y. Sun, "The potential for student performance prediction in small cohorts with minimal available attributes," *Br. J. Educ. Technol.*, vol. 51, no. 2, pp. 347–370, 2020, <https://doi.org/10.1111/bjet.12836>.
- [60] S. Pongpaichet, S. Jankapor, S. Janchai, and T. Tongsanit, "Early detection of at-risk students using machine learning," in *Proc. Int. Conf. Inf. Commun. Technol. Converg. (ICTC)*, pp. 283–287, 2020, <https://doi.org/10.1109/ICTC49870.2020.9289185>.
- [61] A. K. Dileep, A. Bansal, and J. Cunningham, "Early detection of at-risk students in a calculus course," in *Proc. IEEE 46th Annu. Comput., Softw., Appl. Conf. (COMPSAC)*, pp. 187–194, 2022, <https://doi.org/10.1109/COMPSAC54236.2022.00034>.
- [62] A. Pérez-Suay, V. Laparra, S. Van Vaerenbergh, and A. B. Pascual-Ventoe, "Learning about student performance from Moodle logs in a higher education context with Gaussian processes," *IEEE Rev. Iberoam. Tecnol. Aprendizaje*, vol. 19, pp. 169–175, 2024, <https://doi.org/10.1109/RITA.2024.3465035>.