

How Early Can We Identify At-Risk Students? An Analysis of LMS Interactions

Tiloka de Silva and Yomashi Sanjula

University of Moratuwa, Sri Lanka

tilokad@uom.lk

sanjulaagy.21@uom.lk

Abstract: The switch to online learning in higher education brought about by the Covid-19 pandemic has had lingering effects – most notably, continued higher levels of usage of learning management systems (LMS) such as Moodle for assessment and sharing of course materials. This has enhanced the potential for learning analytics even for courses that are delivered in a face-to-face mode. This is because the design of the course page on the LMS and how it is utilized for assessments over the semester necessarily affect the nature of student interactions with the LMS. There is already a sizeable literature that links student interactions with the LMS, selected student characteristics, and learning outcomes, highlighting that it is indeed possible to detect at-risk students using data sources such as course logs and click streams. However, there is less research on how early a student who is at risk of not completing or failing the course can be detected. This paper uses LMS logs, student characteristics, and learning outcomes of six cohorts of undergraduate students (over 500 students in total) taking a compulsory second-year module in a Sri Lankan university to detect the earliest point in the semester at which at-risk students can be identified. Due to the weekly modeling structure, the dataset expands to over 8,000 records, with each entry corresponding to a unique combination of student index number and week number. This paper employed a cumulative modeling approach, where several machine learning models including Random Forest, Decision Tree, Support Vector Machine (SVM), and K-Nearest Neighbors (KNN) are assessed for performance. Random Forest consistently outperformed other models, achieving an accuracy of 78.51% in Week 16. Notably, performance metrics stabilized above 70% by Week 8, suggesting it as the optimal point for early prediction. The analysis revealed that prior academic performance and consistency of LMS engagement were stronger predictors than total LMS clicks. These findings support the development of data-driven early warning systems tailored to the Sri Lankan higher education context, emphasizing the value of consistent behavioral monitoring and historical academic data for effective intervention strategies and it provides insights on how effectively utilizing an LMS can improve learning outcomes even for courses that are offered in face-to-face mode.

Keywords: At-Risk Students, Early Warning system, LMS Interactions, Machine Learning, Learning Analytics

1. Introduction

Learning Management Systems (LMS) are now a core part of modern higher education, revolutionizing how educational institutions deliver and manage learning experiences. The COVID-19 pandemic worldwide significantly accelerated the adoption of the Learning Management Systems (LMS), triggering an unparalleled historical transformation of educational paradigms to e-learning modes. In this context, the necessity for digital and innovative technology to support teaching tasks, manage classes, and track learners has become a critical component in education (Zabolotniaia, 2020; Simanullang and Rajagukguk, 2020). This rapid transition has brought to light opportunities and challenges in terms of digital learning spaces, especially monitoring and supporting the success of learners.

This growth in digital learning spaces has created some unique challenges, particularly in developing countries. In Sri Lanka, for instance, aside from problems relating to access to the suitable infrastructure for online learning (World Bank, 2020), studies have highlighted lack of opportunities for interactive communication between teacher and student, the impossibility of asking a question as soon as some specific issue occurs and the slowdown in development of social skills (Rupasingha and Haththotuwa, 2021). These concerns have highlighted the need to identify students in distance learning environments who are falling behind academically in a timely manner.

In the aftermath of the Covid-19 pandemic, universities resumed normal operations but heightened use of learning management systems for instruction persists (Uthsari et al., 2024). Given that these systems generate vast amounts of behavioral data based on student interactions with course material, assignments, and discussions, they are now being increasingly used to derive insights on learner behaviours (Nguyen, 2015). While much of the literature focuses on the use of LMS data for the study of distance and online learning (e.g. Arizmendi et al., 2022; Kaensar and Wongnin, 2023), there are several papers that investigate outcomes of university students who learn in face-to-face settings too (Uthsari et al., 2024; Biktimirov and Klassen, 2008; Baugher, Varanelli and Weisbord, 2003).

In the case of distance learning courses, the prediction of course completion receives a major focus, given the high levels of non-completion in these courses (Hayes et al., 2024; Latif et al., 2022). Even for courses operating with specific start and end dates (like courses offered in undergraduate or postgraduate degree programs), there is a large literature that connects student learning outcomes with student interactions with the LMS (Kaensar and Wongnin, 2023; Simanullang and Rajagukguk, 2020). Recent advances in educational data mining have leveraged machine learning (ML) approaches to predict student performance using algorithms such as logistic regression, decision trees, support vector machines (SVM), and random forests, which have shown effectiveness in analyzing behavioral patterns and academic histories to support early predictions (Arizmendi et al., 2022; Shayan and van Zaanen, 2019). These methods have the advantage of large and complex datasets, identifying hidden patterns in student behavior and enabling timely interventions to support at-risk students.

Given this established relationship between interactions with the LMS and learning outcomes, another key question arises from the point of view of the educational manager: how early can this data be effectively utilized to identify students at risk of poor academic performance? This is particularly pertinent because even though the use of later-semester data can yield more precise predictions, early identification is the most crucial method of effective intervention to a student. The use of LMS-generated data for addressing this question is also suited for large-class environments where instructors face challenges in closely monitoring individual progress. AI-based tools that analyze early LMS behaviors offer promising solutions by enabling institutions to detect academic risk at earlier stages and initiate support before students fall behind (Latif, 2022).

The existing research on this specific area suggests that the predictive value of different features varies throughout a course. For instance, Shayan and van Zaanen (2019) analyzed data from 426 students across five blended learning courses using Moodle LMS data, student characteristics, and academic performance with a decision tree algorithm and found that prior academic records such as GPA play a significant role in the early weeks, while mid-course performance data and LMS activity metrics (e.g., session counts, clicks, quiz attempts) become more predictive later on. Hayes (2017), in a similar analysis based on LMS behavioural data and predictive modeling finds that the initial six weeks of a course are a critical period in which intervention can have the most impact on student performance. These papers highlight how meaningful insights early in the course are vital for timely interventions though the features used and optimal timing for accurate predictions can vary depending on the context.

Despite advancements in educational data analytics globally, Sri Lankan higher education institutions have yet to effectively utilize LMS behavioral data for early risk prediction. Existing local studies predominantly focus on LMS adoption and student perceptions (Subashini et al., 2022) or examine correlations between LMS usage and academic performance, without offering frameworks for real-time, early-stage interventions (Uthsari et al., 2024; Bandarigodage et al., 2024). Consequently, behavioral indicators available in the initial weeks of a course remain largely untapped for predicting at-risk students. This underutilization limits opportunities for timely academic support and proactive student retention efforts.

Accordingly, this study addresses the following research questions: (i) Can LMS behavioral data and prior academic records predict end-of-semester academic performance in a face-to-face learning environment? (ii) What is the earliest point in the semester at which at-risk students can be identified with acceptable accuracy? (iii) What LMS interaction features most significantly contribute to the prediction of at-risk students? The broader aim of the study is to investigate how predictive models based on early semester data can support the design of effective early intervention strategies within the Sri Lankan higher education context.

To answer these research questions, we analyze LMS interaction data collected during the entire 16 weeks of the semester in a core undergraduate module, together with data on students' end-semester and prior academic performances. A cumulative weekly modeling approach is applied using several machine learning algorithms, incrementally assessing model performance from the first week through the sixteenth week of the semester. Through this design, the research aims to establish an empirical basis for determining the earliest feasible prediction point for academic risk detection, to identify key behavioral indicators associated with performance outcomes, and to propose timely, data-driven intervention strategies that can be operationalized within the Sri Lankan higher education system.

2. Data and Methods

This study investigates the early prediction of academic risk using LMS behavioral data combined with prior academic performance. The analysis workflow includes detailed data collection aligned with weekly course activities, comprehensive data preprocessing to prepare features and address imbalances, and a cumulative

predictive modeling approach that evaluates model performance week-by-week throughout the semester. Each step is designed to address the key research questions concerning the feasibility, timing, and features important for early risk identification in Sri Lankan higher education contexts.

2.1 Data Collection

This study utilizes behavioral and academic performance data from the Faculty of Business, University of Moratuwa. The dataset comprises six cohorts of students (Intakes 17 to 22), focusing on a core module taught in Semester 3, Introduction to Econometrics. The final grade in Introduction to Econometrics serves as the target variable. The grade obtained from its prerequisite subject, Probability and Statistics for Business – II (Semester 2), is included as a predictive variable.

The initial dataset included over 500 unique students. To ensure consistency and avoid duplicate records, students who had repeated attempts in the introduction to Econometrics module were excluded from the analysis. Due to the weekly modeling structure, the dataset expanded to over 8,000 records, with each entry corresponding to a unique combination of student index number and week number. This structure enabled a realistic simulation of ongoing monitoring and prediction throughout the semester.

Behavioral data were extracted from the Moodle Learning Management System (LMS), which records detailed logs of student interactions over the 16-week semester. Weekly clickstream features were derived, including clicks on learning materials, course pages, non-continuous assessment (non-CA) activities, quizzes, and assignments. Each student was identified using an index number, from which the batch prefix was also derived. Week numbers were determined from activity timestamps. These behavioral features were then merged with the final grade and the prerequisite module grade using index numbers as the linking key.

In addition to raw click counts, derived features were engineered to enhance the dataset. These included: (i) the proportion of clicks per activity type within each week, relative to total weekly clicks; and (ii) the week-by-week standard deviation of clicks per activity type, capturing consistency in LMS engagement. All data extraction and feature engineering tasks were performed using Python.

Table 1 provides an overview of the weekly academic activities and the corresponding LMS behavioral features derived from Moodle log files. The weekly breakdown ensures transparency in how student learning behavior was monitored, with LMS interactions captured through event logs (*event name*, *component*, and *event context*). The table is based on the latest course design, with only minor adjustments across different student cohorts.

Table 1: Weekly course structure and LMS interaction mapping

Week	Weekly Activities	LMS Interaction Logs Captured
01	Revision quiz, lecture notes, R code, dataset	Clicks on learning materials, course page, non-CA activities, quizzes
02	Quiz, lecture notes	Clicks on learning materials, course page, quizzes
03	Quiz, lecture notes	Clicks on learning materials, course page, quizzes
04	Lecture notes, R code, dataset	Clicks on learning materials, course page, non-CA activities
05	Lecture notes, R code, dataset, quiz, R activity	Clicks on learning materials, course page, non-CA activities, quizzes
06	Lecture notes, R code, dataset, assignment details	Clicks on learning materials, course page, non-CA activities, assignments
07	Lecture notes, assignment	Clicks on learning materials, course page, assignments
08	Lecture notes, recordings, R code, dataset	Clicks on learning materials, course page, non-CA activities

Week	Weekly Activities	LMS Interaction Logs Captured
09	Lecture notes, R code, dataset, recordings	Clicks on learning materials, course page, non-CA activities
10	In-class test	Clicks on course page, non-CA activities, quizzes
11	Assignment progress presentations	Clicks on course page, non-CA activities, assignments
12	Lecture notes, R code, dataset, recordings, assignment submissions, peer assessments	Clicks on learning materials, course page, non-CA activities, assignments
13	Lecture notes, R code, dataset, quiz, R activity	Clicks on learning materials, course page, non-CA activities, quizzes, assignments
14	R activities	Clicks on course page, non-CA activities, assignments
15	Quiz, Revision questions	Clicks on course page, quizzes
16	Make-up test	Clicks on course page, quizzes

2.2 Data Preprocessing

Several preprocessing steps were performed to prepare the dataset for modeling. Appropriate data types were assigned to each variable, and missing values were handled accordingly. In particular, missing values in the consistency features calculated as the standard deviation of weekly click counts by activity type were observed in Week 1 for all students. This was expected, as standard deviation cannot be computed from a single observation. These values were imputed with zero, indicating the absence of variability in the initial week. Categorical variables were encoded to support downstream modeling. The target variable (final grade) and the prerequisite course grade were ordinally encoded, while batch identifiers were one-hot encoded. All numerical features were standardized to ensure comparability across scales.

Correlation analysis was conducted to identify relationships among features. It provides some interesting insights. Among the LMS interaction variables considered, most are positively correlated with final grade, though the correlations are not strong. The prerequisite course grade was found to be most strongly correlated with the final grade ($r = 0.63$), as shown in Figure 1.

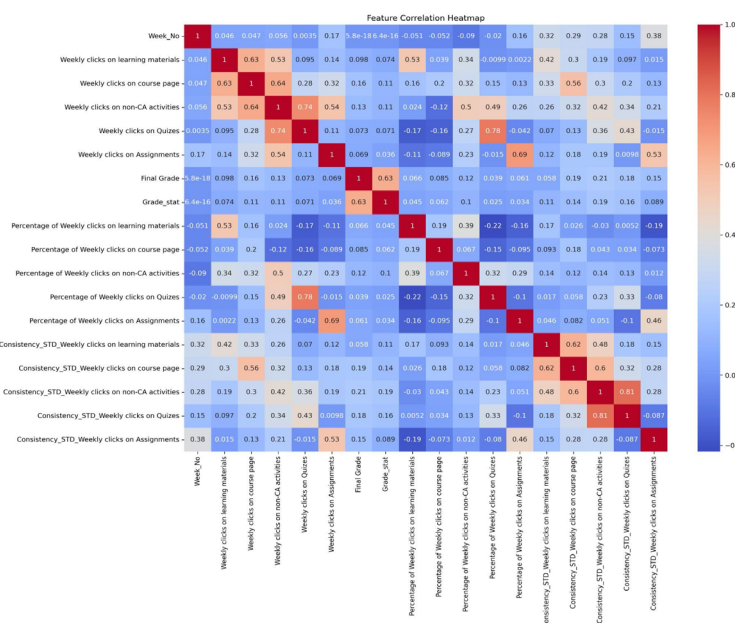


Figure 1: Correlation heatmap of the variables

An analysis of the final grade distribution revealed class imbalance, as shown in Figure 2. To mitigate this, the Synthetic Minority Over-sampling Technique (SMOTE) was applied to balance the dataset. To ensure reliable evaluation of model performance, the dataset was divided into a training set and a testing set, where 80% of the data was used for training and 20% was reserved for testing. This split allows for objectively evaluating how well the models generalize to new unseen data and forms the foundation for the modeling approaches discussed in the next section.

2.3 Modeling Approach

To evaluate the feasibility of early prediction of student performance, a cumulative modeling approach was adopted. Beginning with data from Week 1, models were trained incrementally by adding one week of behavioral data at a time (e.g., Week 1 only, Weeks 1–2, Weeks 1–3, ..., up to Week 16). This enabled performance comparison across different points in the semester and allowed for the identification of the earliest stage at which reliable predictions could be made.

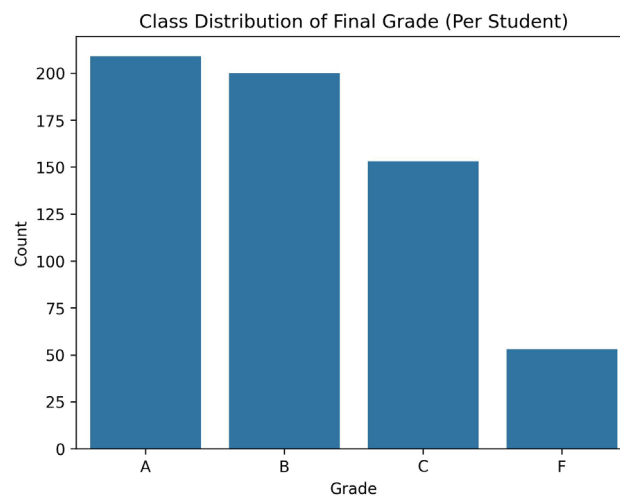


Figure 2: Class imbalance of final grade

The task was framed as a multiclass classification problem, given that the target variable, final grade, is categorical with more than two classes (A, B, C and F). Four classification algorithms were evaluated: Random Forest, Decision Tree, Support Vector Machine (SVM), and K-Nearest Neighbors (KNN). Each model was trained and evaluated on the cumulative datasets using standard multiclass performance metrics, including accuracy, macro-averaged precision, recall, and F1-score. Accuracy measures the proportion of the total number of correct classifications to the sum of classifications. Precision reflects the proportion of students predicted to be at risk who are actually at risk, while recall measures the model's ability to identify all students who are truly at risk. The F1 Score, which is a compromise of both precision and recall, offers a balanced assessment, especially useful when both false positives and false negatives are important to consider. This score is typically an effective metric for measuring the quality of an approach (Shayan and van Zaanen, 2019). The best-performing among these tested models based on these metrics was then selected for further refinement using hyperparameter tuning.

To determine the earliest point of stable prediction, a line graph of model accuracy over time (by week) was plotted. This visualization facilitated the identification of the optimal prediction week—the earliest week in the semester when the model achieved sufficient accuracy to meaningfully differentiate between student performance levels. This point serves as a benchmark for enabling timely and targeted academic interventions.

3. Results

3.1 Model Comparison and Identification of Optimal Week

In the first stage, different models were fitted to the data to predict the final grade and each model was assessed using standard classification metrics to evaluate its performance in predicting at-risk students. The performance metrics across different models are summarized in Table 2.

As Table 2 shows, other than in Week 1, the Random Forest algorithm consistently outperformed other algorithms in classification metrics. Accordingly, to further enhance performance, hyper parameter tuning was applied to the Random Forest model.

Table 2: Comparison of selected models' performance metrics at week 1, week 8 and week 16

Model	Week 1 Performance Metrics	Week 8 Performance Metrics	Week 16 Performance Metrics
Random Forest	Accuracy: 0.5447, Precision: 0.5490, Recall: 0.5447, F1 Score: 0.5359	Accuracy: 0.6972 , Precision: 0.6949 , Recall: 0.6972 , F1 Score: 0.6953	Accuracy: 0.7373 , Precision: 0.7357 , Recall: 0.7373 , F1 Score: 0.7362
Decision Tree	Accuracy: 0.5041, Precision: 0.5067, Recall: 0.5041, F1 Score: 0.5026	Accuracy: 0.6524, Precision: 0.6591, Recall: 0.6524, F1 Score: 0.6549	Accuracy: 0.7154, Precision: 0.7202, Recall: 0.7154, F1 Score: 0.7169
Support Vector Machine (SVM)	Accuracy: 0.5691 , Precision: 0.5891, Recall: 0.5691 , F1 Score: 0.5694	Accuracy: 0.6209, Precision: 0.6285, Recall: 0.6209, F1 Score: 0.6216	Accuracy: 0.6037, Precision: 0.6091, Recall: 0.6037, F1 Score: 0.6056
K-Nearest Neighbors (KNN)	Accuracy: 0.4634, Precision: 0.4929, Recall: 0.4634, F1 Score: 0.4725	Accuracy: 0.5528, Precision: 0.5802, Recall: 0.5528, F1 Score: 0.5625	Accuracy: 0.5574, Precision: 0.5688, Recall: 0.5574, F1 Score: 0.5617

Figure 3 shows the trend of model performance measures throughout the semester of 16 weeks. The model performance shows a clear improvement over time, with the Accuracy (solid blue line) increasing from 0.5366 in Week 1 to 0.7851 in Week 16. While similar improvements were seen in precision (solid orange line) from 0.5490 to 0.7857, recall (solid green line) from 0.5366 to 0.7851, and F1-score (solid red line) from 0.5273 to 0.7846 followed a comparable upward trend, stabilizing around Week 8 and achieving their highest values in the final weeks. This trend in performance is to be anticipated, since each week's model was trained on cumulative behavioral features. That is, more precise LMS interaction data were available later in the semester. As such, the model's ability to distinguish between student final grade categories (A, B, C, F) inevitably improves over time.

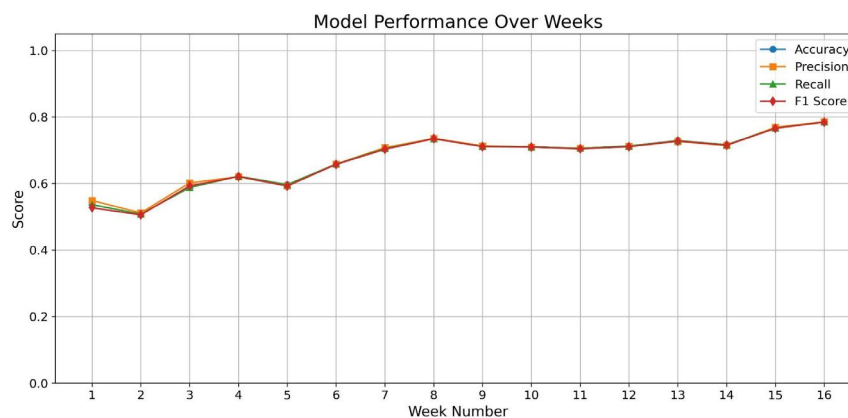


Figure 3: Model performance metrics (Accuracy, Precision, Recall, and F1 Score) across 16 weeks.

Despite relatively modest predictive power during the first few weeks (especially Weeks 1–5), model performance began to increase significantly from Week 5 onward. Most strikingly, by Week 8, all performance metrics exceeded the 0.70 threshold and remained consistently high throughout the remainder of the semester. This stabilization indicates that mid-course, the model has accumulated enough behavioral evidence to yield stable predictions and that Week 8 represents a possible threshold for early detection of at-risk students and intervention planning.

These results confirm the idea that behavioral data extracted from the LMS and aggregated across weeks, is an effective proxy for both academic success and engagement. Yet they also underscore the value of temporal variations showing that while there are some signals in the early weeks, reliable predictions require several weeks of interaction data.

3.2 Feature Importance Analysis

In order to determine which behavioral indicators were most predictive of final grades, feature importance scores were calculated from the best-performing Random Forest model at Week 8, the earliest reliable prediction point identified previously.

As shown in Figure 4, the most valuable feature was `Grade_stat` (importance score: 0.2625), which assesses previous academic performance in the prerequisite statistics course. It accounts for 26.25% of the model performance, whereas LMS interaction variables jointly explain 73.75%. This indicates that although past academic performance is highly influential, behavioral data from the LMS collectively contribute more to the predictive power of the model.

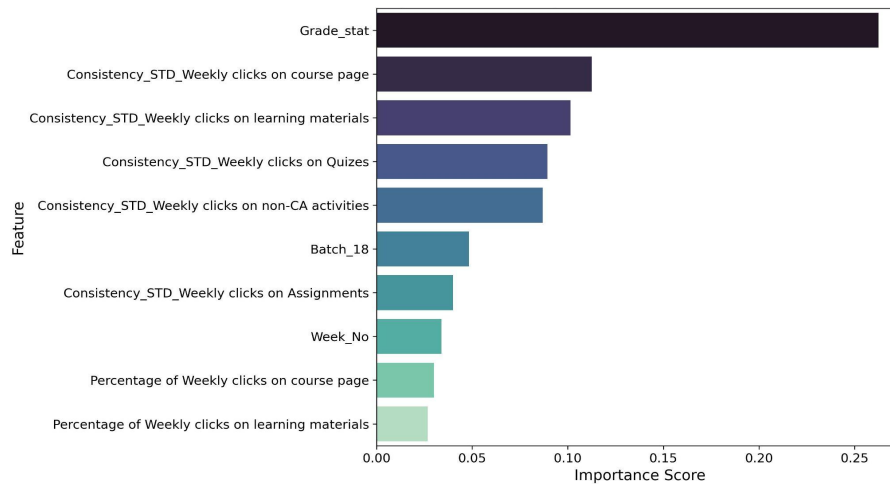


Figure 4: Top 10 most important features identified in Week 8 model for predicting final grades

However, in addition to historical grade information, several features related to consistency in weekly LMS usage have become top-level predictors. These include the standard deviations of weekly clicks on key elements such as course pages, learning materials, quizzes, and non-continuous assessment (non-CA) activities. Their modestly high importance values suggest that students with unstable or highly variable engagement are more likely to be identified as at-risk than students who consistently engage—despite low overall click totals. On the other hand, raw click counts and percentage-based distributions of clicks across activity types showed lower importance, highlighting that sheer volume of activity is less predictive than the pattern or regularity of that activity over time.

These results provide several key implications for early warning systems: first is that prior performance is a key predictor. Second, rather than simply assessing how much a student is utilizing the LMS, it may be more beneficial to track how consistently they use it from week-to-week.

4. Discussion and Conclusions

This study sought to establish how early in a semester behavioral data from an LMS can be used to differentiate between weaker and performing students, with a view to providing timely academic interventions using data from a Sri Lankan university. The findings demonstrate that, aside from prior performance, student usage patterns on Moodle LMS, and particularly those reflecting consistency of engagement, can serve as effective predictors of end-of-term academic performance. These results directly address the first research question, confirming the predictive value of LMS behavioral data combined with prior academic records. The cumulative modeling approach employed in the study demonstrated enhancing model performance measures—accuracy, precision, recall, and F1-score—gradually and consistently with growing quantities of weekly behavior data. Although predictive performance was comparatively modest in the initial weeks, there was a marked improvement starting from Week 5, with the level of performance plateauing above 70% from Week 8. This result indicates that at the mid-point of the semester, there is enough behavioral evidence to effectively distinguish between students who have a likelihood of success and those who may be at risk and it answers the second research question about the timing of at-risk student identification.

The importance of behavioral consistency was further highlighted by feature importance analysis, which revealed that week-over-week fluctuation in LMS activity was more predictive of final outcomes than click

counts. Inconsistent or highly variable interaction patterns were associated with poor performance among students, indicating that participation quality and consistency may be more important than activity counts. In addition, prior academic performance, as captured by the grade in the prerequisite course, emerged as the single most influential predictor of final grade, reaffirming the importance of historical academic data in early risk detection. These findings are consistent with the previous research, which has shown that prior ability strongly influences student performance (Ruiperez-Valiente et al., 2018), and that consistent engagement patterns are more predictive than raw activity volume (Bitkimirov and Klassen, 2008; Baugher, Varanelli and Weisbord, 2003). This analysis provides a clear answer to the third research question by identifying the most significant LMS interaction features contributing to predicting at risk students.

These findings have significant implications for Sri Lankan institutional practices. The identification of Week 8 as a cutoff for stable and accurate prediction offers a handy benchmark for developing early warning systems. Universities can use this insight to develop predictive monitoring tools that automatically flag students showing erratic engagement patterns by mid-semester. Academic support units can deliver focused interventions, like personalized feedback or mentoring, thereby providing students a chance to recover prior to the final assessments.

In conclusion, this study demonstrates that LMS behavioral data, when analyzed cumulatively and combined with prior academic performance, can effectively distinguish between weaker and better-performing students in higher education. Findings confirm that reliable predictions of student outcomes can be made as early as Week 8 of the semester, allowing for a valuable window of opportunity for the early detection of at-risk students. This early prediction capability supports the design of timely, data-driven academic interventions that are both scalable and relevant to the Sri Lankan higher education context. The study further indicates that it is not the quantity of activity in LMS usage, but consistency of usage that more strongly predicts academic performance, and therefore the value of monitoring behavior patterns over time.

This study only addresses one module so is perhaps not generalizable across courses. Though focused on LMS use, instructional approach, course design, and lecturer interaction have the potential to influence learning behaviors that are not quantified directly here. Furthermore, extraneous variables such as economic status or psychological traits were not employed. Future research could extend this analysis to multiple subjects, where LMS usage may vary depending on discipline and instructional methods. A broader study would help assess generalizability and support development of an early intervention framework based on real-time LMS analytics.

Ethics Declaration

This study received ethics approval from the University of Moratuwa Ethics Review Committee under approval number EDN/2023/008. Data collected were de-identified and stored securely. There was no access to individual identifiers such as student names or their contact information at any point in the research process.

AI Declaration

One AI tool (ChatGPT) was used exclusively to aid in enhancing the manuscript's writing quality, grammar, and clarity. The AI tool was not employed to generate content, analyze data, or aid in research design.

References

- Arizmendi, C.J., Bernacki, M.L., Raković, M., Plumley, R.D., Urban, C.J., Panter, A.T., Greene, J.A. and Gates, K.M. (2022). Predicting student outcomes using digital logs of learning behaviors: Review, current standards, and suggestions for future work. *Behavior Research Methods*. doi:<https://doi.org/10.3758/s13428-022-01939-9>.
- Bandarigodage, L., De Silva, T., Ranasinghe, E. and Nanayakkara, V. (2024). Slow and Steady or Fast and Furious: An Analysis of Completion Duration in open.uom.lk. *European Conference on e-Learning*, 23(1), pp.26–35. doi:<https://doi.org/10.34190/ecel.23.1.2667>
- Baugher, D., Varanelli, A., and Weisbord, E., (2003). Student hits in an internet-supported course: How can instructors use them and what do they mean? *Decision Sciences Journal of Innovative Education*, 1(2), 159-179.
- Biktimirov, E.N. and Klassen, K.J., (2008). Relationship between use of online support materials and student performance in an introductory finance course. *Journal of education for business*, 83(3), 153-158.
- De Silva, T., Karunaratne, B., Nanayakkara, V., Karunaratne, B., Ranasinghe, M. and Ranasinghe, E. (2023). e-Learning Interactions and Academic Outcomes: an Analysis of Undergraduates in Sri Lanka. *European Conference on e-Learning*, [online] 22(1), pp.88–96. doi:<https://doi.org/10.34190/ecel.22.1.1629>.

- Haththotuwa, P.M.P.S. and Rupasinghe, R.A.H.M. (2021). Adapting to Online Learning in Higher Education System during the Covid-19 Pandemic: A Case Study of Universities in Sri Lanka. *Sri Lanka Journal of Social Sciences and Humanities*, 1(2), p.147. doi:<https://doi.org/10.4038/sljssh.v1i2.46>.
- Hayes, D., Hong, W., Bernacki, M. and Voorhees, N. (2024). Using LMS Data to Provide Early Alerts to Struggling Students. *Papers on Engineering Education Repository (American Society for Engineering Education)*. doi:<https://doi.org/10.18260/1-2--29442>.
- Kaensar, C. and Wongnin, W. (2023). Analysis and Prediction of Student Performance Based on Moodle Log Data using Machine Learning Techniques. *International Journal of Emerging Technologies in Learning (ijet)*, 18(10), pp.184–203. doi:<https://doi.org/10.3991/ijet.v18i10.35841>.
- Latif, G., Alghazo, R., Pilotti, M.A.E. and Brahim, G.B. (2022). Identifying ‘At-Risk’ Students: An AI-based Prediction Approach. *International Journal of Computing and Digital Systems*, 11(1), pp.1051–1059. doi:<https://doi.org/10.12785/ijcds/110184>.
- Nguyen, T., (2015). The Effectiveness of Online Learning: Beyond No Significant Difference and Future Horizons. *MERLOT Journal of Online Learning and Teaching*, 11(2): 309-319.
- Ruipérez-Valiente, J.A., Muñoz-Merino, P.J., Delgado Kloos, C., (2018). Improving the prediction of learning outcomes in educational platforms including higher level interaction indicators. *Expert Systems*, 35:e12298. <https://doi.org/10.1111/exsy.12298>
- Shayan, P. and van Zaanen, M. (2019). Predicting Student Performance from Their Behavior in Learning Management Systems. *International Journal of Information and Education Technology*, 9(5), pp.337–341. doi:<https://doi.org/10.18178/ijiet.2019.9.5.1223>.
- Simanullang, N.H.S. and Rajagukguk, J. (2020). Learning Management System (LMS) Based On Moodle To Improve Students Learning Activity. *Journal of Physics: Conference Series*, 1462, p.012067. doi:<https://doi.org/10.1088/1742-6596/1462/1/012067>
- Uthsari, S., De Silva, T., Perera, S. and Gunawardana, A. (2024). Online teaching and Learning post-Covid: An Analysis of LMS Usage and Student Outcomes Following the Pandemic. *European Conference on e-Learning*, 23(1), pp.365–373. doi:<https://doi.org/10.34190/ecel.23.1.2746>.
- World Bank, (2020). COVID-19 Response – South Asia: Higher Education [online], World Bank. Available from: <https://documents1.worldbank.org/curated/en/150411590701072157/COVID-19-Impact-on-Tertiary-Education-in-South-Asia.pdf> [Accessed 5 Jun, 2025].