

Article

A Privacy-Conscious AI Framework for Early Identification of At-Risk Students Across Disciplines Using LMS Engagement Data

Hon-Sun Chiu , Adam Wong  and Tung-Lok Wong

Division of Science, Engineering, and Health Studies (SEHS), School of Professional Education and Executive Development, The Hong Kong Polytechnic University, Hong Kong, China; adam.wong@cpce-polyu.edu.hk (A.W.); alan.wong@speed-polyu.edu.hk (T.-L.W.)

* Correspondence: honsun.chiu@cpce-polyu.edu.hk

Abstract

This study explores the application of artificial intelligence (AI) in higher education to enable the early identification of at-risk students using only engagement data from learning management systems (LMS). Unlike many existing early-warning models that are limited to single disciplines or rely on sensitive demographic and prior academic records, the proposed approach offers a privacy-conscious and highly generalizable predictive framework suitable for diverse higher education contexts. The dataset includes over 1.7 million LMS interaction records from 236 undergraduate subjects spanning four academic divisions. These subjects encompass a wide variety of instructional designs and assessment structures. To address cross-subject heterogeneity, this study employs rank-based engagement features that represent students' relative behavioral patterns within each course, facilitating meaningful comparison across disciplines without reliance on absolute activity levels. Using standard machine learning classifiers, the model achieves over 90% prediction accuracy for final subject performance by Week 3 of the semester, demonstrating that reliable early detection of at-risk students is feasible at an early stage of teaching and learning. Rather than claiming intervention effectiveness, the study positions AI-enabled early prediction as a scalable foundation for proactive student support and enhanced teaching responsiveness, with the potential to inform timely pedagogical actions such as targeted outreach and academic advising. By emphasizing generalizability, ethical data use, and privacy protection in AI-enabled learning analytics, this research contributes practical insights into how predictive AI can responsibly support teaching and learning in higher education.

Keywords: artificial intelligence; at-risk students; higher education; learning analytics; early warning systems; educational data mining



Academic Editor: Lan Yang

Received: 2 May 2026

Revised: 25 June 2026

Accepted: 26 June 2026

Published: 1 July 2026

Copyright: © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

1. Introduction

The integration of digital technologies in contemporary higher education has fundamentally transformed the ways in which academic institutions engage with, assess, and support their students. As learning management systems, online assessment platforms, and digital communication tools become ubiquitous, educational institutions can now generate large volumes of student data from multiple sources. This transformation brings both challenges and opportunities for improving student outcomes using advanced analytics. One vital area that has come to the forefront is the early identification of at-risk students. The academic performance, engagement, or well-being of these students suggests a heightened

probability of failing courses, repeating semesters, or discontinuing their studies altogether (Bravo-Agapito et al., 2021; Waheed et al., 2020).

Traditionally, the identification of at-risk students has been a largely manual process, often reliant on educators' experience, observation, and the delayed review of grades or attendance records. Such approaches, while valuable, are inherently limited in their timeliness and objectivity. They may fail to detect subtle or emerging patterns that precede outright academic struggles, and their reactive nature often results in interventions that come too late to effect meaningful change (Li et al., 2023). Recognizing these shortcomings, educational institutions have increasingly turned to artificial intelligence (AI) and machine learning (ML) solutions. These technologies can process large and complex datasets efficiently. They also identify hidden relationships among variables such as online participation, assignment submissions, and engagement patterns (Al-Sulami et al., 2023; Chiu et al., 2023).

The advent of the COVID-19 pandemic has served as a catalyst for the rapid expansion of e-learning and hybrid instruction, fundamentally altering the educational landscape. The abrupt transition to online modalities not only exposed disparities in access and digital literacy but also made it more difficult for educators to monitor student engagement and well-being through traditional means (Bozkurt & Sharma, 2020; Murphy, 2020). In this context, the role of AI in educational analytics has become more prominent. AI-driven predictive models can seamlessly integrate a variety of data types, including quantitative metrics like grades, qualitative measures such as sentiment analysis from discussion posts, and demographic backgrounds, to forecast student risk with greater accuracy and enable earlier, more proactive interventions (Kumar et al., 2025; Luo, 2024).

These predictive analytics tools can deliver actionable insights directly to instructors, counselors, or academic advisors. These insights highlight students who may benefit from additional support, personalized feedback, or counseling services (Russell et al., 2020). By doing so, AI-based systems have the potential to inform timely pedagogical responses that support academic engagement and retention, and foster a more inclusive and responsive educational environment. Furthermore, AI enables continuous monitoring and real-time feedback. These capabilities can help to create a culture of data-driven decision-making, where institutional strategies are refined based on empirical evidence rather than anecdotal observations.

However, the promise of AI in identifying at-risk students does not come without its own set of challenges. Integrating and standardizing data from disparate sources remains a technical hurdle, especially in institutions employing varied digital ecosystems. Student privacy and ethical considerations must be rigorously maintained, with transparent policies to safeguard personally identifiable information and ensure compliance with legal frameworks. Issues of algorithmic bias, equity, and model interpretability are also critical, as poorly designed systems may inadvertently reinforce existing disparities or deliver opaque recommendations that are difficult for stakeholders to trust or act upon (Hanna et al., 2025; Ujkani et al., 2024).

Given this complex backdrop, the purpose of this research article is to systematically explore the use of AI in identifying at-risk students within higher education. By focusing on generalizable, privacy-conscious early risk detection, our study aims to lay a foundation for enhancing teaching responsiveness and proactive student support, both of which are core concerns in higher education analytics. Using only student engagement data from the learning management system (LMS), our prediction model does not access sensitive personal information and achieves over 90% accuracy in forecasting students' final subject performance by the third week of the semester. This study makes three main contributions:

- (i) First, it proposes a privacy-conscious AI framework that relies solely on LMS engagement data, thereby avoiding the use of demographic and prior academic information.
- (ii) Second, it introduces rank-based engagement features as a cross-disciplinary standardization mechanism that enables meaningful comparison across heterogeneous course designs.
- (iii) Third, it demonstrates that early prediction of academic risk is feasible by Week 3 of the semester across a large and diverse dataset comprising multiple disciplines.

We begin with a comprehensive literature review that synthesizes the current state of research on AI-driven risk identification. Then, the methodology section outlines the design and implementation of our AI prediction within a higher education setting, followed by a presentation of key findings. In the discussion, we analyze the implications of our results and propose recommendations for effective and ethical deployment of AI-based systems. Finally, the conclusion summarizes our contributions and offers directions for future research in this rapidly evolving field.

2. Literature Review

Several studies have highlighted the importance of early detection and intervention. These studies emphasize the need for comprehensive frameworks that integrate AI techniques to analyze student data effectively (Bravo-Agapito et al., 2021; Waheed et al., 2020; Li et al., 2023). The findings suggest that implementing AI-driven predictive models can significantly enhance educators' ability to identify and support at-risk students, ultimately improving their academic success and well-being (Al-Sulami et al., 2023; Kumar et al., 2025).

Learning analytics and data-driven approaches play a crucial role in identifying and addressing discrepancies between coursework and exam performance among at-risk students (Ifenthaler & Yau, 2020; Russell et al., 2020). These methodologies leverage big data and machine learning to predict student outcomes, identify at-risk individuals, and provide personalized interventions (Bravo-Agapito et al., 2021; Chiu et al., 2023). By analyzing various data sources, such as learning management systems (LMS), previous academic performance, and socio-demographic information, institutions can proactively support students who may struggle academically (Lakkaraju et al., 2015; Sneyers & De Witte, 2018). This approach not only helps in early identification but also in tailoring educational strategies to meet individual needs, thereby enhancing student success and retention (Alyahyan & Dustegor, 2020).

2.1. Traditional Approaches to Identifying At-Risk Students

Conventional methods for identifying at-risk students have relied predominantly on static, lagging indicators such as cumulative grade point averages (GPA), exam scores, and attendance rates (Waheed et al., 2020; Li et al., 2023; Ifenthaler & Yau, 2020). These techniques have been criticized for their reactive nature and limited ability to capture the dynamic, multifaceted processes underlying student success or failure (Alyahyan & Dustegor, 2020). Even though some studies acknowledged the value of social and demographic variables, such as socioeconomic status and first-generation status, these data points often failed to provide a holistic picture of student risk (Lakkaraju et al., 2015; Sneyers & De Witte, 2018). Manual identification processes were also susceptible to inconsistencies and subjectivity, potentially overlooking subtle signs of disengagement or distress (Li et al., 2023). As a result, interventions were frequently implemented too late to effect meaningful change (Waheed et al., 2020).

Increasingly, the complexity of student populations, coupled with the vast data generated by LMS, digital assessments, and online participation, has made AI an indispensable

asset (Li et al., 2023). AI-driven models can integrate and analyze a broad spectrum of data sources, including student engagement logs, discussion forum participation, socio-demographic profiles, and even sentiment analysis from written submissions, offering a more holistic and proactive approach (Alyahyan & Dustegor, 2020).

2.2. AI and Machine Learning Approaches on Predicting At-Risk Students

The COVID-19 pandemic served as a catalyst for the accelerated adoption and ongoing refinement of AI-driven educational analytics. During this period, universities and colleges were confronted with unprecedented requirements for remote monitoring and student support (Bozkurt & Sharma, 2020; Murphy, 2020; Daniel, 2020; Means & Neisler, 2021). In response to the abrupt shift away from face-to-face interaction, there was a notable surge in the implementation of digital learning platforms and analytics tools (Bozkurt & Sharma, 2020). These tools became essential for maintaining instructional continuity and monitoring student engagement. Daniel (2020) highlighted that the pandemic also exposed significant disparities in digital access, with marginalized students being disproportionately affected. This underscored the critical need for adaptive, equity-focused analytics that could help identify and support vulnerable learners. Meanwhile, Means and Neisler (2021) emphasized the importance of AI-enabled early warning systems, which became crucial for detecting signs of student disengagement and emerging mental health concerns among those learning remotely. The period witnessed a proliferation of research studies assessing the effectiveness of predictive models, particularly in online and blended learning environments (Wang et al., 2024). These insights remain relevant in post-pandemic blended and digitally mediated learning environments.

The emergence of AI and ML has enabled sophisticated early warning systems. These systems utilize real-time and multidimensional data from multiple sources. Modern predictive models now integrate behavioral patterns obtained from LMS, social network analysis, assignment submission timings, forum activity, and even biometric data (Bravo-Agapito et al., 2021; Chiu et al., 2025; Alyahyan & Dustegor, 2020). A systematic review conducted by Bravo-Agapito et al. (2021) underscored the effectiveness of ensemble machine learning techniques, which combine multiple models and data sources to enhance predictive accuracy. Similarly, Lakkaraju et al. (2015) suggested that ensemble and deep learning models, particularly when trained on longitudinal data, can identify at-risk students with greater precision than traditional statistical methods.

Early prediction using LMS data has been explored in prior studies, demonstrating its feasibility across different learning contexts. Early work by Hu et al. (2014) showed that time-dependent variables extracted from LMS records are effective in predicting online learning performance. Their study also indicated that learner performance can be characterized while the course is still in progress. Later studies, such as Akçapınar et al. (2019a) and Foster and Siddle (2020), also highlighted the value of behavioral data for identifying at-risk students at an early stage. Akçapınar et al. (2019a) showed that students with unsuccessful end-of-term outcomes could be identified early using online interaction data. Similarly, Foster and Siddle (2020) reported that no-engagement alerts derived from learning analytics data were more efficient than demographic data in identifying students at risk of poorer outcomes. More recent research has continued to investigate early prediction in diverse contexts, including machine learning-based prediction and system-level implementations (Luo, 2024; Wang et al., 2024; Chong et al., 2025; Dai et al., 2025).

Luo (2024) explored the integration of sentiment analysis and clickstream data into risk prediction models. Their findings indicated that social-emotional indicators, such as changes in student communication tone or participation, often precede academic decline. Alyahyan and Dustegor (2020) reinforced that models drawing on both cognitive

and non-cognitive variables, including motivation and self-regulation, provided a fuller understanding of risk trajectories. Furthermore, [Ifenthaler and Yau \(2020\)](#) discussed the growing importance of dashboard visualizations for communicating predictive analytics insights to instructors and advisors in an actionable format. Similarly, [Bañeres et al. \(2020\)](#) developed and tested an early warning system in online higher education. Their system allowed stakeholders to analyze risk information through dashboards and supported the use of early feedback as an intervention mechanism.

The role of intervention has also received sustained attention in the literature. [B. T. M. Wong and Li \(2020\)](#) reviewed 24 case studies of learning analytics interventions in higher education. Their review showed that intervention practices commonly focused on improving study performance, offering personalized feedback, and supporting retention. It also highlighted challenges related to scalability, implementation conditions, delivery channels, and evaluation of intervention effectiveness. More recent work has further integrated prediction with intervention design. [Bañeres et al. \(2023\)](#) proposed a system that identifies dropout risk and supports intervention at the assessable-activity level. Their results showed that interventions aimed at goal setting on ongoing assessable activities reduced dropout issues and increased engagement. [Dai et al. \(2025\)](#) also combined early identification with feedback intervention. Their study showed that predictive models can be used to identify at-risk students in a new cohort and that feedback intervention can improve subsequent behavioral engagement.

A number of studies have explored a range of datasets, feature sets, and ML algorithms to improve the early identification and intervention for at-risk students. For example, [Adnan et al. \(2021\)](#) focused on identifying at-risk students in virtual learning environments at the earliest possible stage, using the Open University Learning Analytics Dataset (OULAD). They applied several algorithms and found that Random Forest models showed progressive improvements in accuracy as more course data became available. [Alturki et al. \(2022\)](#) targeted the prediction of academic status and grades among master's students, employing demographic and post-enrollment features and achieving high accuracy (77–94%) with Random Forest models. Similarly, [Alwarthan et al. \(2022\)](#) addressed early risk detection in preparatory year students at Imam Abdulrahman bin Faisal University, using demographic, pre-university, and assessment data. Random Forest model again outperformed others, with accuracies between 91.7% and 99.7% across multiple datasets.

Other studies, such as [Gonzalez-Nucamendi et al. \(2023\)](#), explored dropout prediction among first-year undergraduates by integrating demographic, financial, and admission data, finding that their Random Forest model was particularly effective in identifying retained students (99.4%). [Jamjoom et al. \(2021\)](#) examined at-risk prediction in a programming course, with both Decision Tree and Support Vector Machine models achieving 99.18% accuracy. [Jang et al. \(2022\)](#) compared a range of algorithms on online course data and found that Logistic Regression offered the best overall performance in terms of AUC and F1 measures, although ensemble approaches sometimes excelled in specific weeks. [Karalar et al. \(2021\)](#) investigated ensemble models during the pandemic, demonstrating that combinations of Extra Trees, Random Forest, and Logistic Regression achieved a specificity of 90.34%. [T. Liu et al. \(2022\)](#) introduced the GDPN model, a hybrid of Gate Recurrent Unit Neural Network and Deep Neural Network, which achieved a 92.5% accuracy rate in predicting student performance using the OULAD. These research studies are summarized in Table 1.

Table 1. Summary of existing research on predicting at-risk students.

Research Studies	Research Purpose	Dataset	Features	Algorithm with Best Performance Achieved	Scope
Adnan et al. (2021)	Early identification of at-risk students in virtual learning environments using pass/fail binary classification	Open University Learning Analytics Dataset (OULAD)	Demographics, Clickstream, Assessment	Random Forest Average accuracy = 59%, 79%, 84%, 88%, 90%, 91% at 0%, 20%, 40%, 60%, 80% and 100% of course length	Online study mode Not clearly specified: Discipline; With/without exam
Alturki et al. (2022)	Predicting master's students' academic status (completed or not) and grade (above average, average, below average)	Over 700 student records from 2010–2018 at the Faculty of Business Informatics and Mathematics, University of Mannheim, Germany	Demographics, Post-enrollment features	Random Forest Accuracy: 77–94%	Single discipline; With exam; Not clearly specified: Study mode
Alwarthan et al. (2022)	Predict at-risk students in the preparatory year at an early stage using binary classification (At-risk and Not At-risk)	Three datasets for humanities preparatory year students at Imam Abdulrahman bin Faisal University (IAU)	Demographic data, Pre-University data, Courses grades, Cumulative GPA, Assessments marks	Random Forest Accuracy– Dataset 1: 99.662% Dataset 2: 91.715% Dataset 3: 95.363%	Single discipline; With exam; Not clearly specified: Study mode
Gonzalez-Nucamendi et al. (2023)	Predicting the first-year undergraduate student dropouts	Dataset contains first-year undergraduate student of Tecnológico de Monterrey	Demographic data, Pre-admission data, Admission data, Financing data	Random Forest Precision– Total sample: 92% Dropped out: 13.2% Retained: 99.4%	Not clearly specified: Discipline; Study mode; With/without exam
Jamjoom et al. (2021)	Using binary classification (Yes: passed; No: not passed) to predict at-risk students in a programming language course	Dataset consists of assessments and grading of the course at the Computer Science Department of Princess Nourah bint Abdulrahman University	Pre-admission scores, Assessments marks	Decision Tree & Support Vector Machine Accuracy: 99.18%	Single course; With exam Not clearly specified: Study mode
Jang et al. (2022)	Predict at-risk students using binary classification	Data from five undergraduate courses and two graduate courses delivered online at Korea University during the academic year 2020–2021	Demographics, Online behavior, Assessment Score	Logistic Regression has the best overall performance in terms of AUC measures and F1 measures (except Stacking Classifier in Week 2 and Random Forest in Week 3)	Online study mode; With exam Not clearly specified: Discipline

Table 1. Cont.

Research Studies	Research Purpose	Dataset	Features	Algorithm with Best Performance Achieved	Scope
Karalar et al. (2021)	Predicting students at risk of academic failure during the pandemic	Data of a 15-week compulsory course of all first-grade students by the end of the 2020-fall semester in a Turkish state university	Degree, Gender, Moodle Activity; Adobe Connect	Ensemble model consists of combinations of Extra Trees, Random Forest and Logistic Regression Specificity: 90.34%	Single course; With exam Not clearly specified: Study mode
T. Liu et al. (2022)	Predict student performance using binary classification (Pass and Fail)	Data related to the FFF course in the Open University Learning Analysis Data Set (OULAD)	Clickstream	Proposed GDPN Accuracy: 92.5%	Single course; Online study mode Not clearly specified: With/without exam

2.3. Key Predictors and Model Effectiveness

Building on the representative studies summarized in Table 1, the broader literature indicates substantial diversity in data sources, feature selection strategies, and model choices. Among these studies, demographic, academic, and behavioral variables are most commonly employed as predictors, either in raw form or through varying degrees of feature engineering, depending on dataset availability and institutional context. Despite the growing maturity of predictive analytics in education, there remains no universally accepted framework that standardizes which datasets, variables, or supervised learning models are most appropriate for reliably and consistently identifying at-risk students.

For instance, studies utilizing large-scale learning analytics datasets, such as the OULAD, which contains demographic and behavioral variables, have reported differing conclusions regarding both model performance and key predictors. Al-Shabandar et al. (2019) found that Gradient Boosting Machines (GBM) achieved the highest predictive accuracy. They also found that behavioral engagement indicators were the most influential predictors in at-risk identification. These indicators included time- and clickstream-based variables. Similarly, Ujkani et al. (2024) suggested that a custom neural network outperformed other models on OULAD data. They identified “total_clicks” as the most significant feature. In contrast, Bainbridge et al. (2015) applied demographic, academic, and behavioral data from an online Master of Public Administration program at Marist College using a Logistic Regression (LR) model. They reported that two of the academic performance variables exerted the strongest influence, whereas demographic factors contributed minimally to predicting students’ at-risk status. Related work has also pointed to the importance of combining multiple data sources when predicting student outcomes. Vaarma and Li (2024) used transcript, demographic, and LMS data to predict student dropouts in Finnish higher education. Their results showed that LMS activity should be considered alongside transcript and demographic variables. They also showed that feature importance and prediction performance can vary over time.

Other studies illustrate further variation in model effectiveness and predictor importance. Gray et al. (2016) analyzed demographic and examination data from first-year students across multiple disciplines at the Institute of Technology Blanchardstown, Ireland. They compared six base classification models and two ensemble classification models. Their results showed that k-nearest neighbors (kNN) achieved the highest accuracy. They also found that demographic variables alone had higher predictive power in predicting students’

final performance in the new-year cohort. Helal et al. (2018) compared interpretable “white box” approaches with black-box classification models using first-year local students’ enrolment and LMS activity data of a division in an Australian university. They reported that no single model consistently outperformed others across all evaluation metrics in predicting at-risk students. Their findings highlighted the advantage of combining enrolment and behavioral LMS features, while also emphasizing the value of model interpretability in educational contexts. Karalar et al. (2021) discovered that an ensemble model combining Extra Trees (ET), Random Forest, and LR produced high predictive accuracy using LMS and virtual classroom activity data of a Turkish state university. Their results showed that gender was among the least influential predictors relative to behavioral variables in predicting students at risk of academic failure. Likewise, Nimy et al. (2023) proposed a probabilistic logistic regression model that comprises students’ demographic, academic, and LMS behavioral data of a South African university. They specified “TIME ON COURSES” as the most salient predictor of academic risk.

Differences also emerge across study modes and analytical scope. While several studies explicitly focused on fully online learning environments (Bainbridge et al., 2015; Al-Shabandar et al., 2019; Karalar et al., 2021; Ujkani et al., 2024), others either implicitly targeted face-to-face settings or did not specify modality (Gray et al., 2016; Helal et al., 2018; Nimy et al., 2023). Similarly, some investigations spanned multiple disciplines, whereas others analyzed data from a single course or module. Despite these variations, many studies highlight the predictive value of behavioral LMS data, particularly time-based engagement and clickstream indicators in predicting at-risk students.

Several studies further addressed the temporal dimension of early warning systems, suggesting that meaningful prediction is possible at relatively early stages of a course. Nimy et al. (2023) claimed week 6 as the optimal point for accurately identifying the at-risk students, while Helal et al. (2018) reported that “vulnerable” students could be detected between weeks 4 and 6, with certain at-risk profiles identifiable even before course commencement. Gray et al. (2016) similarly argued that risk prediction could occur prior to the start of the first year. Collectively, these findings underscore both the promise of AI-driven early prediction and the ongoing challenge of reconciling methodological diversity, contextual variability, and interpretability in predictive learning analytics.

In addition to individual empirical studies, recent review papers provide a broader perspective on current developments in the field. Chong et al. (2025) reviewed current trends in the use of machine learning to predict student engagement and academic performance in online higher education. Their review emphasized the growing importance of this line of research for educational planning and early identification in online learning environments. Topali et al. (2025) examined human-centered learning analytics and artificial intelligence in education. Their review showed that many studies involve stakeholders in the design of solutions, but fewer studies involve them during ideation and prototyping. The review also noted that many studies do not report evaluation. These findings suggest that the design, adoption, and evaluation of AI-based educational systems remain important concerns in the field.

2.4. Challenges and Ethical Considerations

While AI offers significant promise, the literature has identified several challenges and ethical dilemmas associated with its use in educational contexts. A significant technical challenge in building effective at-risk student prediction models is the pervasive issue of class imbalance (Thammasiri et al., 2014). In most educational datasets, students labeled as “at-risk” represent only a small fraction of the overall student population, resulting in models that can become biased toward predicting the majority class, typically those

students who are not at risk. To overcome this, various approaches have been adopted. The Synthetic Minority Over-sampling Technique (SMOTE), for example, creates synthetic samples of the minority class by interpolating between existing at-risk cases, thus enhancing the diversity and representation of at-risk students within the training data (Chawla et al., 2002). Compared to other methods for addressing class imbalance, such as random oversampling or undersampling, SMOTE offers distinct advantages. Random oversampling simply duplicates minority class samples, which can lead to overfitting by memorizing specific cases rather than generalizing patterns. Undersampling, meanwhile, removes samples from the majority class, potentially discarding valuable information.

The issue of data imbalance can also be addressed using sklearn. The resample method in sklearn is designed for general resampling purposes, such as upsampling or downsampling datasets, and can also be utilized for bootstrapping. This approach has been applied in image analytics research (A. Wong et al., 2024). Nonetheless, it is less advanced compared to SMOTE. SMOTE takes a more sophisticated approach by generating synthetic minority instances through interpolation, which creates new, plausible data points rather than mere copies. This helps models learn a broader representation of at-risk students' behaviors without sacrificing existing data or introducing excessive redundancy (Waheed et al., 2020; Bravo-Agapito et al., 2021). Furthermore, SMOTE's ability to create a diverse set of samples reduces the risk of overfitting and enhances the model's sensitivity to subtle risk indicators, leading to fairer and more robust early warning systems (Ifenthaler & Yau, 2020; Waheed et al., 2020). However, although SMOTE has been used in higher education settings (A. Wong et al., 2023), synthetic oversampling remains limited by both the absolute scarcity and sometimes the homogeneity of "at-risk" cases, especially in the early weeks of a class when few students have demonstrated clear signs of risk (Chawla et al., 2002). This makes early intervention particularly challenging from a data perspective.

Equally important are the ethical issues inherent to AI-driven student risk prediction. Privacy is paramount, as educational data can include sensitive personal information ranging from assessment records to behavioral logs. Institutions must rigorously enforce data protection practices and comply with regulations, which mandate transparency, data minimization, and secure handling of personal information (Hanna et al., 2025). Algorithmic fairness is also a central concern: AI models must be regularly audited for potential biases that could arise from historical inequities in the data or from the disproportionate influence of certain features (Ding et al., 2019). For instance, over-reliance on demographic or socioeconomic features can inadvertently perpetuate systemic disadvantages for under-represented populations. It should be highlighted that most recent research studies, as listed in Table 1, have a strong reliance on demographic, academic, and behavioral data, where ethical and privacy issues are concerned.

Recent research has explored formal privacy-preserving approaches in educational data analytics. Federated learning has been applied in learning analytics to train models across distributed datasets without centralizing sensitive student data, thereby reducing privacy risks while maintaining predictive performance (Fachola et al., 2023). In addition, differential privacy has been proposed as a framework to protect individual-level data in learning analytics by introducing controlled noise into model training processes, balancing privacy protection and analytical accuracy (Q. Liu et al., 2025). Another similar work has also integrated differential privacy with interpretable machine learning models for student performance prediction, showing that privacy guarantees can be achieved with only modest reductions in predictive accuracy (Zhao et al., 2025). These approaches provide stronger technical privacy guarantees compared with traditional data anonymization or data-minimization strategies. However, they often require additional infrastructure, parameter tuning, and validation across institutional contexts.

In contrast, the present study adopts a data-minimization approach by using only LMS engagement data and excluding demographic and historical academic variables. While this approach does not provide formal privacy guarantees, it offers a practical and scalable alternative for institutions seeking to balance predictive analytics with privacy considerations.

Another challenge is generalizability. AI models developed and trained by data from a particular course, institution, or population often underperform when applied to a new setting (Tsiakmaki et al., 2020). This lack of transferability stems from substantial differences in curricular structures, teaching methods, assessment types, grading policies, and the backgrounds of student cohorts. For example, a model that accurately predicts student risk in a data-rich, STEM-focused course at a large research university may struggle when deployed in a small liberal arts class or in institutions serving non-traditional student populations (Kumar et al., 2025). In these contexts, features that are highly predictive in one environment, such as frequency of online discussion participation or specific quiz scores, may be irrelevant or even misleading elsewhere. As a result, extensive re-training, domain adaptation, or even customized model development becomes necessary to achieve acceptable performance, increasing the workload and technical demand of educational technologists and data scientists.

Most studies rely on raw or aggregated engagement data. These data are often combined with demographic and academic variables. Hu et al. (2014), Foster and Siddle (2020), and Akçapınar et al. (2019a, 2019b) all demonstrate the predictive value of LMS-based behavioral indicators. However, the literature reviewed above places much less emphasis on transforming behavioral data into relative rankings within peer groups in the same subject. This gap provides the rationale for the approach adopted in the present study. The proposed approach focuses on privacy-conscious design and cross-disciplinary generalizability.

For instance, Akçapınar et al. (2019b) have carried out a study that tested the effects of transformed variables in predictive models. The study compared the prediction results of 3 types of datasets. These included a raw dataset, a transformed dataset that used percentile rank transformation to convert raw data into values between 0 and 1, and a categorical dataset that discretized the transformed data. The results showed that the transformed dataset produced the lowest predictive performance compared with the raw and categorical datasets. Dien et al. (2020) have conducted an analysis that assessed the feasibility of data transformation. A technique called Quantile Transformation (QTF) was used to convert all the features to a range of 0 to 1 and input them into the prediction models. They found that when data transformation is utilized, the models yield better prediction outcomes. Although both studies used a similar data transformation approach that converted the features ranging from 0 to 1, Akçapınar et al. (2019b) focused on the eBook interaction logs of a single course, while Dien et al. (2020) made use of the demographic and academic data of multiple disciplines.

In response to these challenges and building on the above studies, our research is focused on developing a generalized AI prediction model that is applicable across a wide range of disciplines and subject areas. Importantly, our approach adopts a data-minimization strategy to reduce reliance on sensitive data and promote ethical use. To this end, the model relies exclusively on course-related data and learning behavioral data from LMS log files, which are readily accessible to instructors for each course. While LMS engagement data is behavioral and potentially identifiable, this study adopts a data minimization approach by excluding demographic and historical academic variables. In addition, we define rank-based engagement features as variables obtained by ranking each student's behavioral activity relative to peers within the same subject. These features capture relative engagement patterns rather than absolute activity levels, enabling meaningful cross-subject comparison.

3. Methodology

This research is situated within the School of Professional Education and Executive Development (SPEED) at The Hong Kong Polytechnic University (PolyU), an institution recognized for its commitment to academic excellence in professional and postgraduate education. PolyU SPEED offers a diverse array of top-up degree and taught postgraduate programmes, encompassing disciplines organized across four distinct divisions, namely the Division of Business and Hospitality Management (BHM), the Division of Languages and Communication (LC), the Division of Science, Engineering and Health Studies (SEHS), and the Division of Social Sciences, Humanities and Design (SSHD).

The academic calendar is structured around two primary semesters per year, each comprising 13 weeks dedicated to teaching and learning activities. The first two weeks constitute the add/drop period, during which enrolment may fluctuate. A formal census point in Week 3 subsequently finalizes enrolment. These administrative timelines can influence early engagement patterns and should be taken into account when interpreting predictive performance in the initial weeks. Assessment structures vary across subjects, typically including 3 to 5 components such as assignments, tests, or projects. In general, the first assignment is scheduled between Weeks 3 and 5, followed by a mid-term test between Weeks 7 and 9. The second assignment or group project is typically due between Weeks 11 and 13. For some subjects without an examination component, a final test is usually conducted in the last teaching week. These were not explicitly modeled due to data granularity limitations, but may represent useful features for future work.

The LMS utilized by PolyU SPEED is Blackboard, a comprehensive digital platform designed to facilitate course administration, content dissemination, and student engagement. Blackboard enables instructors to construct individualized course sites, wherein students are granted access to a broad spectrum of educational resources, including lecture materials and assignments. The system further supports interactive pedagogical practices by providing integrated tools for online discussions, assignment submissions, and real-time quizzes.

3.1. Data Collection

Data collection spanned the second semester of the 2023–2024 academic year, encompassing 236 top-up degree subjects offered by four academic divisions. Comprehensive subject grades were collected for all students across each course. Certain subjects include examination components, while others do not. As shown in Figure 1, the distribution of grades approximates a normal curve, with the mean falling within the B range. For this study, at-risk students are defined as those whose grades are C– or lower, representing approximately 5% of the dataset.

Additionally, student behavioural data relating to learning activities were extracted from Blackboard into a student activity log file with over 1.7 million rows of records. Table 2 lists the columns of the log file. Students may enrol in several subjects during a semester and may demonstrate varying performance across these subjects. To accurately capture learning habits in each subject, the combination of PersonUsername and CourseNumber serves as a unique identifier for each record. Such uniquely defined records enable robust tracking and analysis of each student's engagement with course materials, offering granular insights into learning behaviours across multiple subjects.

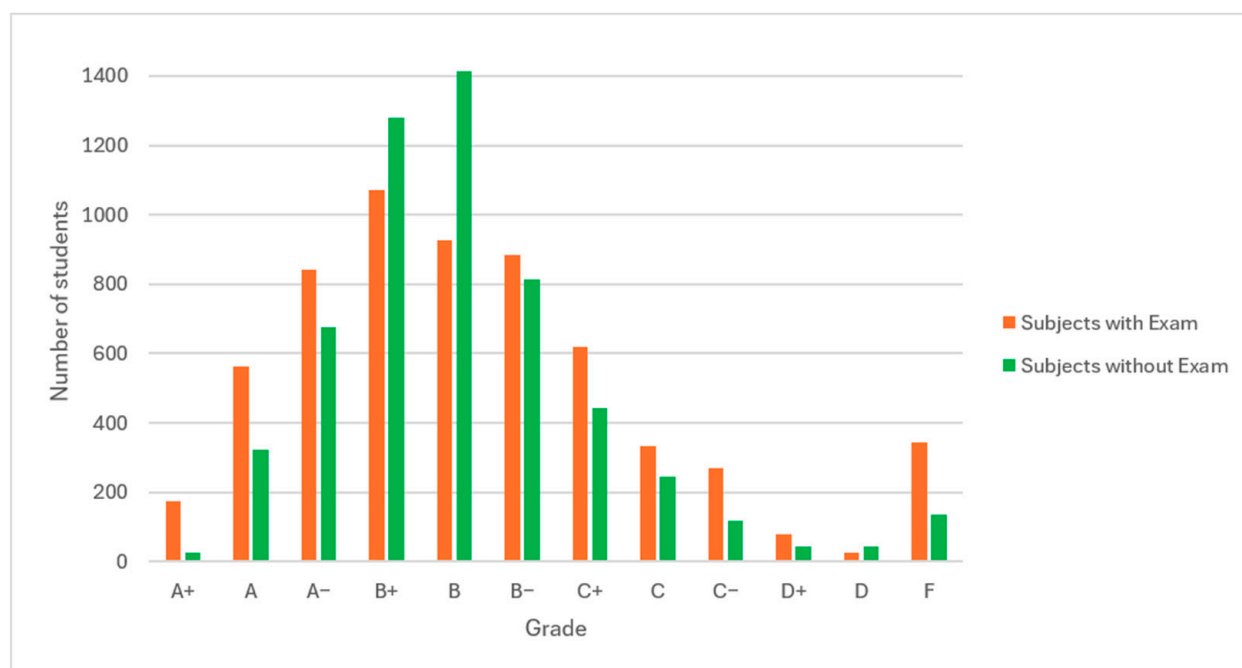


Figure 1. Distribution of student grades across all subjects in the 2023–2024 semester 2.

Table 2. Columns of student activity log file.

Column	Definition
TimeAccessed	The time stamp that tells when the students access Blackboard.
Date	The date that the students access the content in Blackboard.
PersonUsername	The student ID number of the students.
CourseNumber	The code of the top-up degree subject.
CourseItem	The items that the students accessed.
CourseItemType	The category that the items (CourseItem) belong to.
TotalTime	The total time that the students spend on the items (CourseItem). The time unit is minutes.
TotalClicks	The total number of clicks on the items (CourseItem).
SessionID	The ID for each session. Records in the same session have the same ID.
TimeInSession	The sum of TotalTime in the same session (counted per course). The time unit is minutes. The value is the same for records in the same session.
ClicksInSession	The sum of TotalClicks in the same session (counted per course). The value is the same for records in the same session.

3.2. Data Preprocessing

Based on the student activity log data extracted from Blackboard, we applied feature engineering techniques to construct new variables that encapsulate students' weekly learning patterns across the 13 teaching weeks of the semester. The DATE attribute in the log file was utilized to determine each teaching week. Recognizing that learning is a cumulative process, we performed cumulative data aggregation for each period: Week 1, Weeks 1–2, Weeks 1–3, and so forth, continuing through Weeks 1–13. The definition of new features is listed in Table 3.

Table 3. New features created by feature engineering.

Feature	Definition
Week	The period of cumulative data from week 1 to the current week.
Days	The number of days the student accessed Blackboard during the specified period.
Sessions	The number of sessions (identified by SessionID) the student accessed Blackboard during the specified period.
Total_Time	The total time (the sum of TotalTime) the student spent in Blackboard during the specified period.
Avg_Total_Time	The average time (the average of TotalTime) the student spent accessing the items in Blackboard during the specified period.
Std_Total_Time	The standard deviation of time (the standard deviation of TotalTime) the student spent accessing the items in Blackboard during the specified period.
Total_Clicks	The total number of clicks (the sum of TotalClicks) recorded by the student in Blackboard during the specified period.
Avg_Total_Clicks	The average number of clicks (the average of TotalClicks) recorded by the student on the items in Blackboard during the specified period.
Std_Total_Clicks	The standard deviation of clicks (the standard deviation of TotalClicks) recorded by the student on the items in Blackboard during the specified period.
Avg_Time_Per_Day	The average amount of time the student spent accessing Blackboard per day during the specified period (calculated by $\text{Total_Time}/\text{Days}$).
Avg_Clicks_Per_Day	The average number of clicks recorded by the student in Blackboard per day during the specified period (calculated by $\text{Total_Clicks}/\text{Days}$).
Avg_Time_Per_Session	The average amount of time the student spent accessing Blackboard per session during the specified period (calculated by $\text{Total_Time}/\text{Sessions}$).
Avg_Clicks_Per_Session	The average number of clicks recorded by the student in Blackboard per session during the specified period (calculated by $\text{Total_Clicks}/\text{Sessions}$).

To assess the suitability of these engagement features for predicting academic performance, we conducted a correlation analysis between these features and the subject grades across the four divisions. The results of this analysis are presented in Table 4. It is important to note that the correlations between digital engagement features and subject grades observed in our study are notably low, which is much lower than what is commonly reported in the literature, where studies often consider a single subject and find moderate to strong relationships. This discrepancy can be attributed primarily to the scale and diversity of our research. We analyzed data across 236 distinct subjects, each with its own instructional design, assessment structure, and learning activities. The requirements and expectations for student engagement on Blackboard vary significantly from one subject to another. For instance, spending 10 hours on Blackboard might be excessive for a course with a heavy emphasis on in-person activities, but could be insufficient for a fully online subject requiring extensive digital participation. As a result, aggregate measures of engagement, such as time

spent or number of clicks, lose their predictive power when generalized across a highly heterogeneous set of subjects. This nuanced landscape makes it challenging to interpret raw engagement metrics as universally “good” or “bad” indicators of academic performance.

Table 4. Correlation between features and subject grades among 4 divisions.

	BHM	SEHS	LC	SSHD
Days	0.203	0.185	0.142	0.175
Sessions	0.169	0.160	0.116	0.158
Total_Time	0.075	0.123	0.189	0.114
Avg_Total_Time	0.090	0.089	0.154	0.122
Std_Total_Time	0.100	0.118	0.200	0.114
Total_Clicks	−0.064	0.077	0.088	0.030
Avg_Total_Clicks	−0.075	0.036	−0.076	0.077
Std_Total_Clicks	−0.071	0.052	−0.042	0.055
Avg_Time_Per_Day	0.008	0.020	0.108	0.034
Avg_Clicks_Per_Day	−0.213	−0.125	−0.056	−0.192
Avg_Time_Per_Session	−0.008	0.003	0.099	0.018
Avg_Clicks_Per_Session	−0.231	−0.151	−0.081	−0.221

To enhance the generalizability of features for predicting outcomes across various subjects, we observed that students with higher academic performance tend to exhibit greater engagement with Blackboard. Figure 2 presents the distribution of student activity on Blackboard by grade, highlighting distinct engagement patterns. The analysis indicates a consistent trend, where students earning higher grades regularly demonstrate increased participation on Blackboard, as evidenced by more frequent logins, extended time spent, and a greater number of interactions with course materials. These findings suggest a strong association between active engagement on the platform and improved academic achievement.

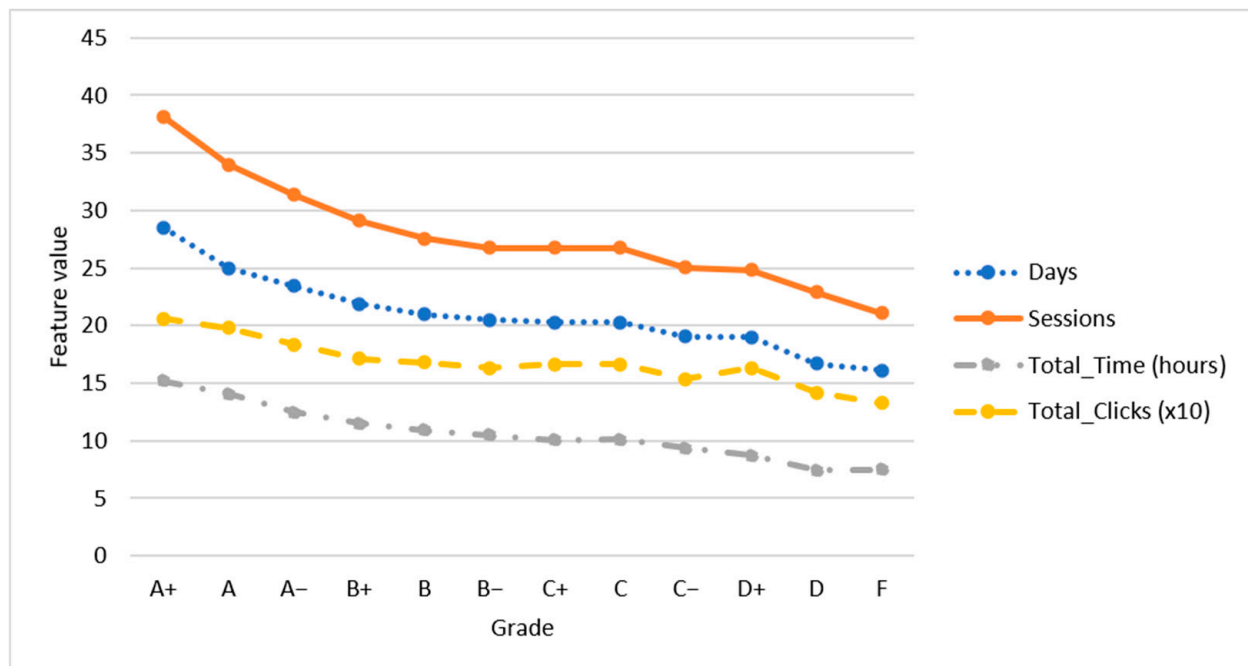


Figure 2. Blackboard activity of students by grades.

Accordingly, rather than utilizing the raw features listed in Table 4, we transformed students’ LMS engagement measures into rank-based features by ranking students’ behav-

ioral activity within each subject. This approach captures relative engagement rather than absolute activity levels, thereby reducing the influence of subject-specific differences and enabling more meaningful cross-subject comparison. Specifically, Rank_Days, Rank_Sessions, Rank_Total_Time, and Rank_Total_Clicks, were derived from Days, Sessions, Total_Time, and Total_Clicks, respectively. These rank-based features not only enrich the input for model prediction but also standardize engagement measures across all subjects. As previously noted, significant variation in Days, Sessions, Total_Time, and Total_Clicks exists between different subjects, making direct comparison difficult. The use of rank-based features, therefore, provides a more consistent basis for comparing student engagement across heterogeneous teaching and learning contexts. An additional benefit of the rank-based features is their ability to maintain a comprehensive engagement record for each student, including those who have not participated in any activities on Blackboard. Traditionally, raw features lacking a value (e.g., when a student does not login to Blackboard during the first week) were excluded from use in the prediction model. By utilizing rank-based features, however, these students are now assigned the lowest rank in the subject. This approach ensures that their entries remain in the dataset, thereby offering additional input to enhance the predictive model.

It is important to recognize that the Rank_Total_Time (derived from Total_Time) feature may not accurately represent student engagement, as students might simply open an item on Blackboard and leave it active without further interaction. Using this metric alone can be misleading. Consequently, a new feature called Index has been developed, defined as the sum of Rank_Days, Rank_Sessions, and Rank_Total_Clicks. The Index feature provides a more comprehensive measure of student engagement. Table 5 summarizes these rank-based features.

Table 5. Rank-based features.

Feature	Definition
Rank_Days	The rank of the student in a subject in terms of number of days the student accessed Blackboard during the specified period.
Rank_Sessions	The rank of the student in a subject in terms of number of sessions the student accessed Blackboard during the specified period.
Rank_Total_Time	The rank of the student in a subject in terms of total time the student spent on Blackboard during the specified period.
Rank_Total_Clicks	The rank of the student in a subject in terms of total number of clicks recorded by the student on Blackboard during the specified period.
Index	The sum of Rank_Days, Rank_Sessions and Rank_Total_Clicks.

3.3. Handling Imbalanced Dataset

The dataset is divided into 3 groups for classification: Well-performing (Grade A and above); Normal (Grade C to A−); and At-risk (Grade C− and below). Including the well-performing class in predictive modeling offers substantial benefits beyond the conventional focus on at-risk students. By identifying students who are excelling, institutions gain a comprehensive perspective on the full spectrum of academic engagement and achievement. This allows educators to recognize effective learning behaviors and strategies present in top performers, which can inform the development of support programs, peer mentoring, and best practices that benefit the broader student population. Moreover, understanding the attributes of well-performing students helps in setting realistic benchmarks and aspirational

goals for other learners, fostering a culture of continuous improvement. Ultimately, a balanced approach that includes both ends of the performance spectrum enables more targeted, equitable, and holistic educational strategies.

However, as shown in Figure 1, well-performing students and at-risk students are the minority groups of the dataset. As mentioned above, this kind of class imbalance is a common challenge in educational data mining and predictive analytics, particularly for early warning systems designed to identify students at risk of academic failure (Bravo-Agapito et al., 2021; Waheed et al., 2020). To mitigate the issue of class imbalance, we adopted SMOTE to generate synthetic samples for the minority classes. To avoid data leakage, SMOTE was applied only to the training data after the train-test split. The dataset was first divided into training and testing sets, after which SMOTE was applied to the training set to balance class distributions. This ensures that synthetic samples do not influence model evaluation on unseen data. The procedure was repeated on a weekly basis to ensure that the training set remains balanced for each week's predictions. By artificially enlarging the representation of well-performing and at-risk students in the training set, SMOTE helps to balance the class distribution, allowing predictive models to better recognize patterns associated with different groups rather than simply favoring the majority class.

4. Experimental Results

For the prediction of students' subject performance using AI, both regression and classification techniques were utilized. The regression analysis employed Random Forest Regressor and Multiple Linear Regression models, while the classification task was addressed using the Random Forest Classifier.

The choice of predictive models in this research reflects both the complexity of educational data and the need for robust, interpretable results. Random Forest Regressor was selected for regression tasks due to its ability to handle non-linear relationships, interactions among variables, and resistance to overfitting. These are common challenges in educational datasets characterized by diverse student behaviors and engagement metrics. This ensemble method aggregates predictions from multiple decision trees, improving accuracy and generalizability even when predictor variables are correlated or noisy (Lakkaraju et al., 2015).

Multiple Linear Regression, on the other hand, was employed for its interpretability and effectiveness when relationships between predictors and outcomes are largely linear. In educational research, linear models provide clear insights into the contribution of each feature (such as engagement indices) to student performance, supporting actionable decision-making for educators and administrators (Li et al., 2023). Moreover, its widespread use in early educational analytics allows for benchmarking against established findings in the literature (Ifenthaler & Yau, 2020).

For classification tasks, the Random Forest Classifier stands out due to its superior performance in differentiating between multiple categorical outcomes, such as the three student performance groups defined in this study. Its ensemble structure makes it particularly robust to imbalances and noise in the data, a significant advantage when dealing with minority classes like well-performing and at-risk students (Bravo-Agapito et al., 2021; Chiu et al., 2023). Additionally, Random Forest Classifier can reveal important patterns and feature importances that are crucial for early warning systems in educational settings (Lakkaraju et al., 2015; Ujkani et al., 2024).

4.1. Feature Selection Using SelectKBest

Feature selection analysis was conducted on all variables presented in Tables 3 and 5 using the SelectKBest method. SelectKBest is a feature selection method that enables the identification of the most relevant predictors for a given task by ranking all available features according to a specified statistical test and then retaining only the top K features with the highest scores (A. Wong et al., 2023). This approach improves model performance by reducing dimensionality, mitigating overfitting, and enhancing interpretability, as it ensures that only the most informative variables are included in the predictive modelling process (Ifenthaler & Yau, 2020).

Feature selection is performed weekly for each subject. Due to the extensive volume of results, only a summary can be presented here. Figure 3 provides a screenshot of SelectKBest output of the first three weeks of a specific subject as an example, where the top ten features are indicated above the red line and rank-based features are highlighted in yellow. It is generally observed that rank-based features receive higher selection scores. To illustrate this selection preference, the weekly average scores for both rank-based and non-rank features are calculated, and these averages are subsequently used to determine their ratio of selection preference by the system.

Subject	Number of Students	Week 1		Week 1 - Week 2		Week 1 - Week 3	
		Feature	Score	Feature	Score	Feature	Score
SEHS4595	189	Avg_Clicks_Per_Day	67.492	Rank_Total_Clicks	35.587	Rank_Total_Clicks	36.437
		Rank_Total_Clicks	64.978	Index	33.759	Total_Clicks	36.426
		Index	51.172	Rank_Sessions	27.228	Avg_Clicks_Per_Day	35.271
		Total_Clicks	39.790	Total_Clicks	26.661	Index	34.444
		Sessions	24.387	Sessions	26.398	Rank_Sessions	30.378
		Rank_Sessions	24.387	Avg_Clicks_Per_Day	18.237	Sessions	29.558
		Rank_Total_Time	15.811	Rank_Total_Time	14.065	Total_Time	17.353
		Avg_Clicks_Per_Session	12.780	Days	12.653	Days	10.293
		Rank_Days	6.168	Rank_Days	12.653	Rank_Days	10.293
		Days	6.168	Total_Time	10.408	Rank_Total_Time	9.911
		Std_Total_Clicks	1.770	Avg_Total_Clicks	8.894	Avg_Total_Clicks	7.622
		Total_Time	0.667	Avg_Time_Per_Day	3.738	Avg_Clicks_Per_Session	5.782
		Avg_Time_Per_Day	0.594	Avg_Clicks_Per_Session	3.441	Std_Total_Clicks	4.940
		Std_Total_Time	0.593	Std_Total_Clicks	2.910	Avg_Total_Time	3.731
		Avg_Total_Time	0.592	Std_Total_Time	0.599	Avg_Time_Per_Day	2.178
		Avg_Time_Per_Session	0.255	Avg_Time_Per_Session	0.424	Std_Total_Time	0.207
		Avg_Total_Clicks	0.176	Avg_Total_Time	0.000	Avg_Time_Per_Session	0.143

Figure 3. A screen capture of SelectKBest results.

Figure 4 presents the selection ratio between rank-based and non-rank features as determined by SelectKBest. The week number on the x-axis represents data included cumulatively from Week 1 through that specific week. The results indicate a preference for rank-based features, attributed to their generalized representation of student engagement across various subjects. Furthermore, rank-based features consist of consecutive values, which enhance their linear correlation with subject grades. In contrast, non-rank features tend to exhibit greater variability and lack linearity, making them less effective as predictors of subject grades. Based on this rationale, rank-based features are utilized in subsequent predictive analyses.

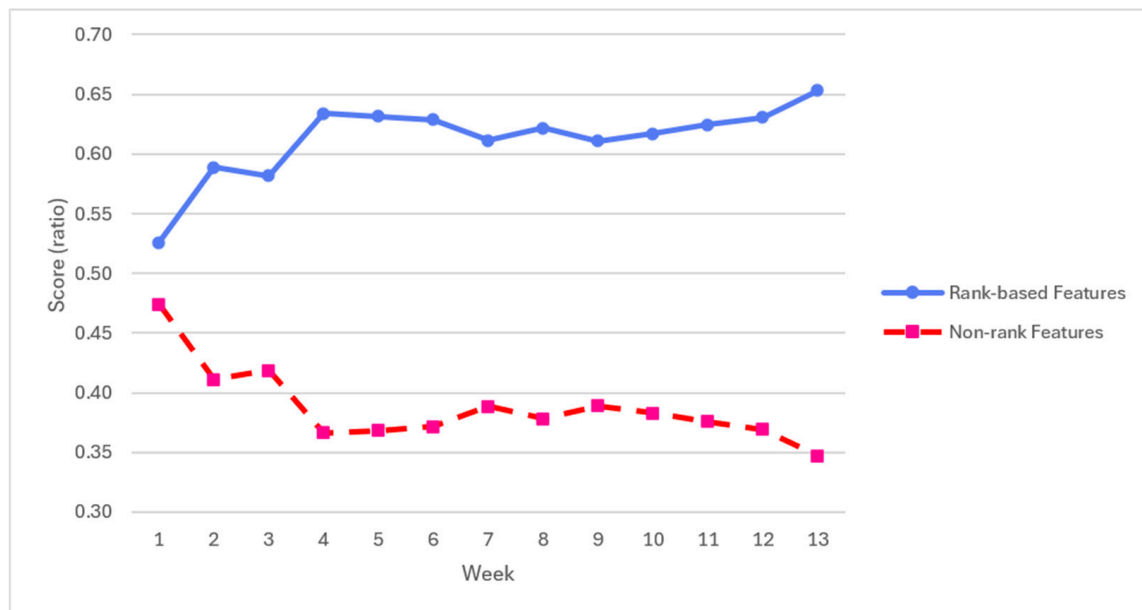


Figure 4. Selection ratio of rank-based and non-rank features by SelectKBest.

4.2. Baseline Prediction Result of Large Subjects

To begin with, we assessed the predictive performance using data from large subjects that included an examination component. Performance prediction is especially crucial in the context of large subjects, where class sizes often exceed 80 students, and individualized attention becomes increasingly challenging. In these expansive learning environments, instructors face significant obstacles in closely monitoring the academic progress, engagement, and well-being of each learner. Traditional methods such as manual grade checks, classroom observations, or one-on-one meetings are often impractical at scale, and students at risk of falling behind may go unnoticed until it is too late to implement effective support.

A total of 16 subjects met the large subject criteria, encompassing 2341 students. The relevant data were extracted from the dataset and divided into training and testing sets using a stratified 70/30 split at the student record level. This ensured that all data associated with a given student were confined entirely to one set. A stratified 5-fold cross-validation scheme was then applied within the training partition to guide model tuning. Subsequently, the training sets were balanced using SMOTE. The results are reported across cumulative weekly datasets to provide robustness across temporal slices.

The prediction performance of the regression approaches is measured by using the R-squared (R^2) and Mean Squared Error (MSE), which are widely adopted as evaluation metrics for regression models due to their complementary strengths in interpreting predictive accuracy. While R^2 reflects overall explanatory capacity, MSE ensures that individual prediction errors remain minimal. Together, they provide a balanced view of both model fit and prediction precision.

For the classification approach, accuracy and F1 score are used as evaluation metrics. Accuracy measures the overall proportion of correctly classified instances out of all predictions, offering a straightforward indication of general model effectiveness. However, accuracy alone can be misleading, as a model could achieve high accuracy by favouring the majority class and overlooking minority groups. The F1 score addresses this limitation by harmoniously balancing precision (the proportion of relevant instances among the retrieved instances) and recall (the proportion of relevant instances that were retrieved). By considering both false positives and false negatives, the F1 score ensures that the model

does not sacrifice sensitivity to the needs of underrepresented groups for the sake of overall correctness.

Figures 5–7 present the prediction performance of the three approaches. It is observed that Random Forest Regressor (Figure 5) outperforms Multi-linear Regression (Figure 6) with both higher R2 values and lower MSE. For the Random Forest Classifier (Figure 7), both accuracy and F1 score achieve above 80%. It should be noted that regression and classification models address different prediction tasks and are therefore not directly comparable. The regression models are presented for exploratory purposes, while the main focus of this study is on classification performance.

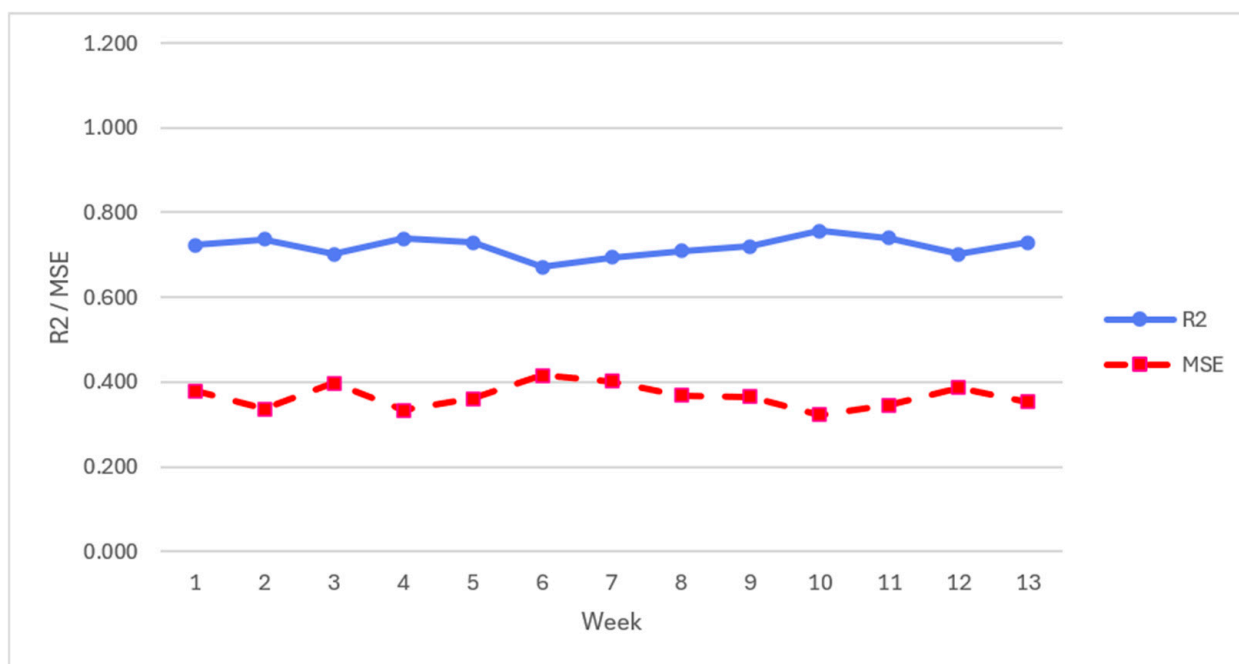


Figure 5. Prediction performance for large subjects using Random Forest Regressor.

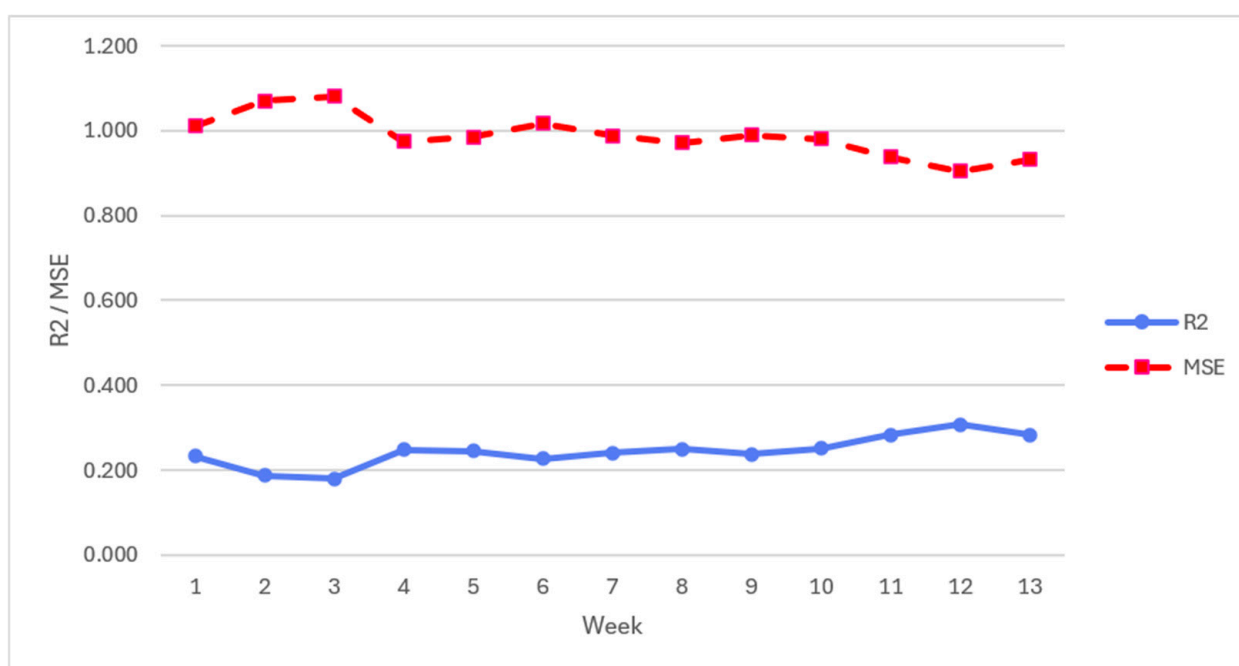


Figure 6. Prediction performance for large subjects using Multiple Linear Regression.

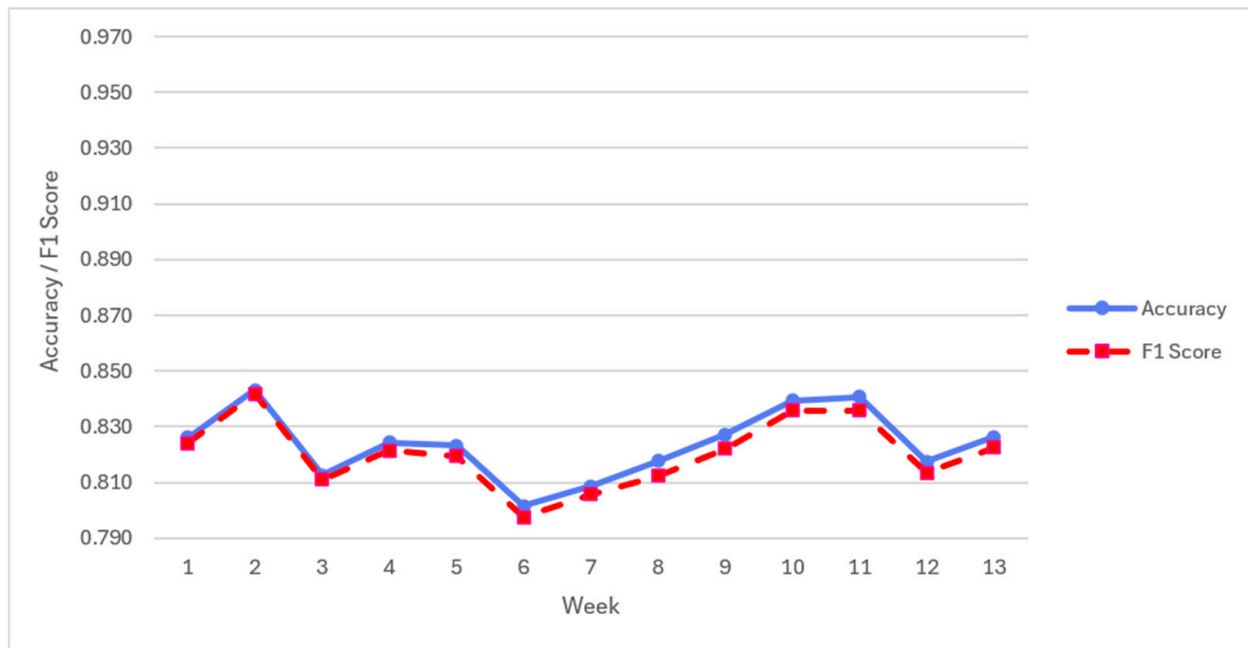


Figure 7. Prediction performance for large subjects using Random Forest Classifier.

4.3. Prediction of All Subjects Using Classification

Analysis of the baseline prediction results indicates that classification using rank-based features yields more than 80% of prediction accuracy of subject grades. Consequently, a Random Forest Classifier is employed to predict classifications for all students across 236 subjects. However, several subject-specific variables may influence predictive performance when utilizing a generalized model. To address this, 3 additional features are incorporated into the prediction inputs:

- EXAM (indicating whether the subject includes an examination)
- MODE (the instructional mode: face-to-face, online, or hybrid)
- DIVISION (the academic division associated with the subject)

In combination with the 5 rank-based features detailed in Table 5, these 8 input features collectively offer a more comprehensive representation of students' learning characteristics under different subject settings.

Figures 8–11 present the prediction results for all subjects using a classification approach with the Random Forest Classifier, segmented by class size. In particular, when comparing the prediction performance of large subjects in Figures 7 and 8, it is observed that incorporating the 3 additional subject-specific features results in a notable increase in prediction accuracy from 80% to over 90%. The analysis also indicates that prediction performance is higher for large subjects compared to medium and small subjects in Figures 9 and 10. This may be attributed to the larger student populations in these subjects, where grade distributions are more likely to approximate normality. Conversely, medium and small subjects exhibit less distinct grade distribution patterns, resulting in a more complex relationship between rank-based features and subject grades.

Another observation is that accuracy and F1 score are very close in value in all plots. This is due to the fact that the model has successfully mitigated majority-class bias after applying SMOTE to the training data. Furthermore, the overall error distribution demonstrates that both false positive and false negative rates remain proportionally low across the classification space. It is also observed that both accuracy and F1 scores are lower during the initial two weeks of the semester. This can be attributed to the add/drop period in the first 2 weeks, during which subject enrolment has not yet been finalized. As students may

enter or withdraw from subjects in this timeframe, there is often insufficient Blackboard engagement data available for reliable subject grade prediction.

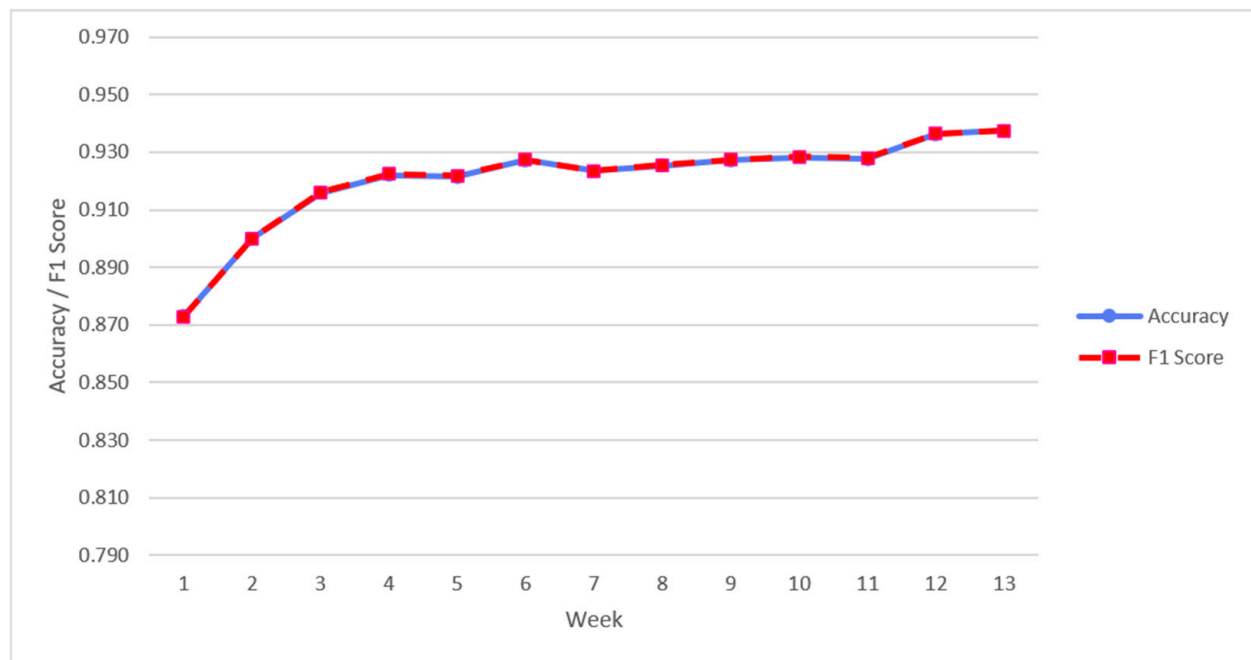


Figure 8. Prediction performance using classification for large subjects (class size above 80).

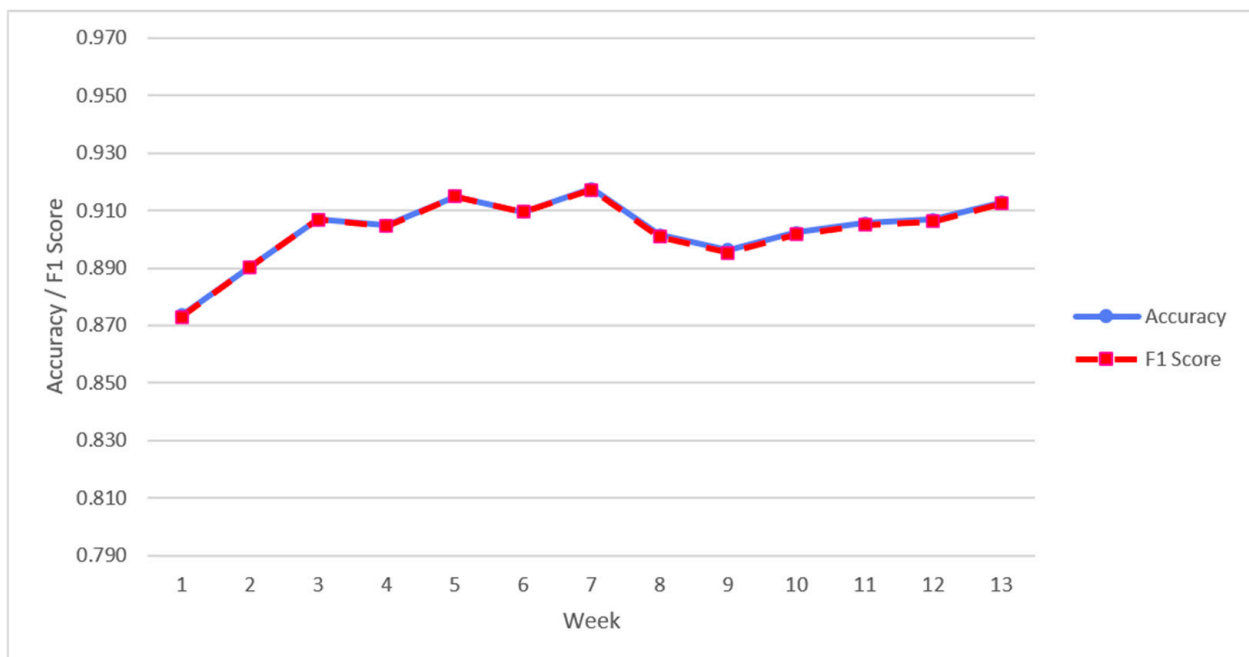


Figure 9. Prediction performance using classification for medium subjects (class size 50–80).

Figures 12–14 present the per-class evaluation results for the classification model across all 236 subjects, including precision, recall, and F1 score. Interestingly, these per-class metrics reveal that the minority classes (At-risk and Well-performing) outperform the majority class (Normal) in both precision (in Figure 12) and F1 score (in Figure 14), while maintaining comparable recall (in Figure 13). This phenomenon indicates that balancing the training dataset with SMOTE enabled the model to establish highly precise decision boundaries for the minority classes. As a result, the model effectively minimized

false positive rates for the At-risk and Well-performing classes without reducing overall sensitivity. This explains why their precision exceeded that of the inherently noisier majority class.

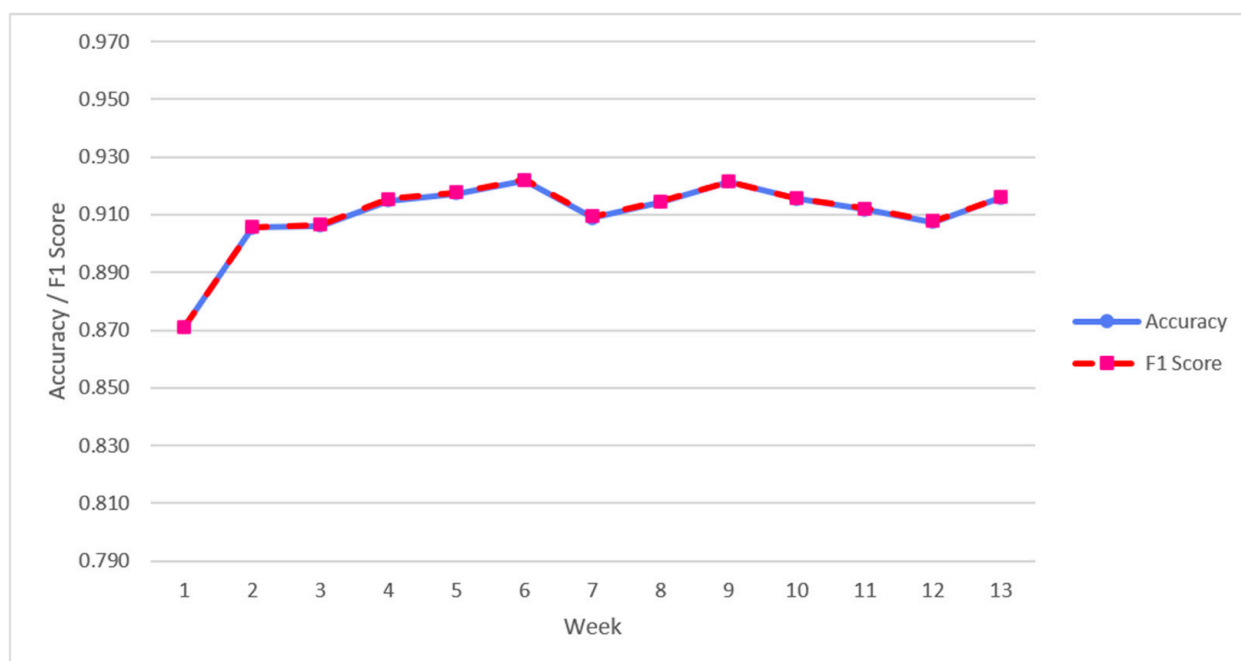


Figure 10. Prediction performance using classification for small subjects (class size below 50).

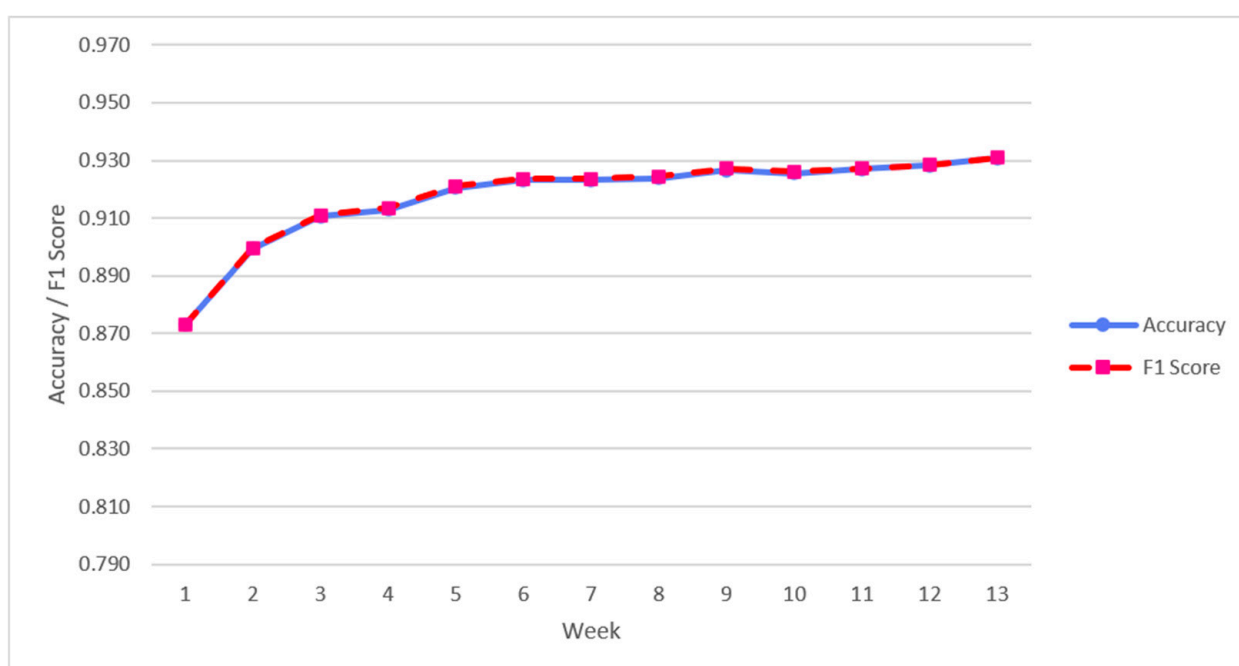


Figure 11. Prediction performance using classification for all 236 subjects.

The overall performance of the predictive models is highly encouraging. Notably, both the prediction accuracy and the F1 score surpass 90% from the third teaching week onwards. These findings reflect the model's robust capacity to correctly classify student outcomes and minimize misclassifications. This level of predictive power early in the semester is particularly significant, as it demonstrates that meaningful, actionable insights can be derived well before final assessments or grades are determined. The results provide

empirical evidence supporting the feasibility of early prediction and its potential value for instructional support. The results also affirm that actionable insights can be generated as early as the third teaching week, with substantial implications for student success and institutional effectiveness.

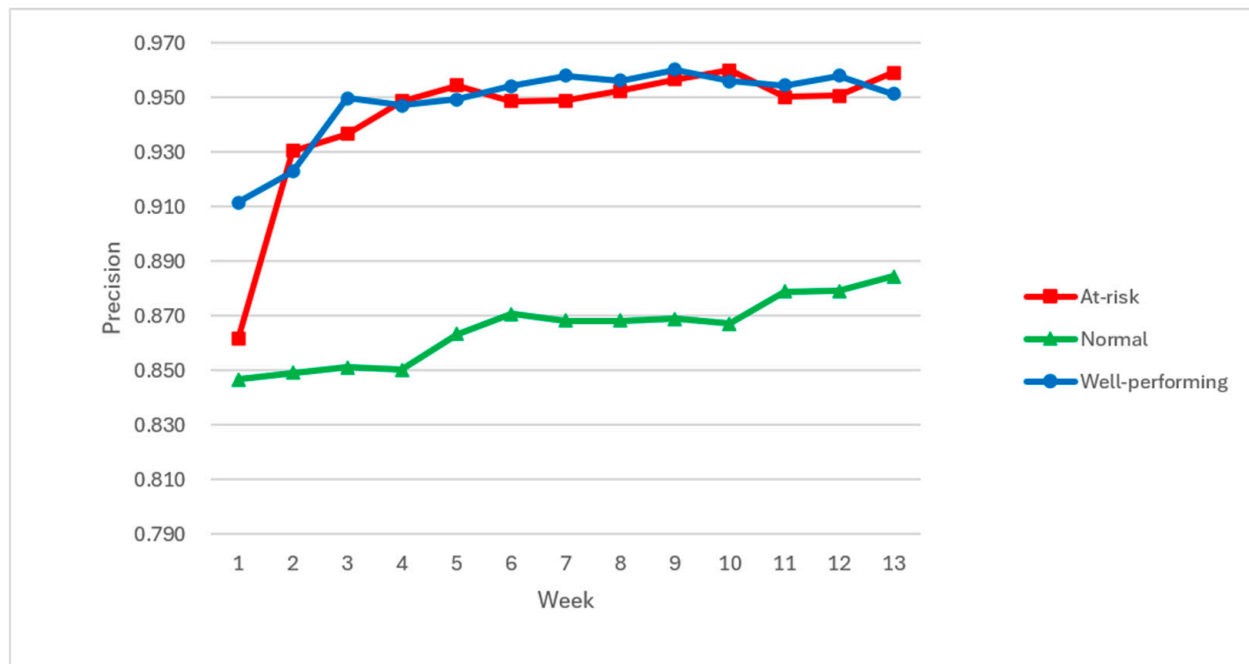


Figure 12. Per-class precision of classification for all 236 subjects.

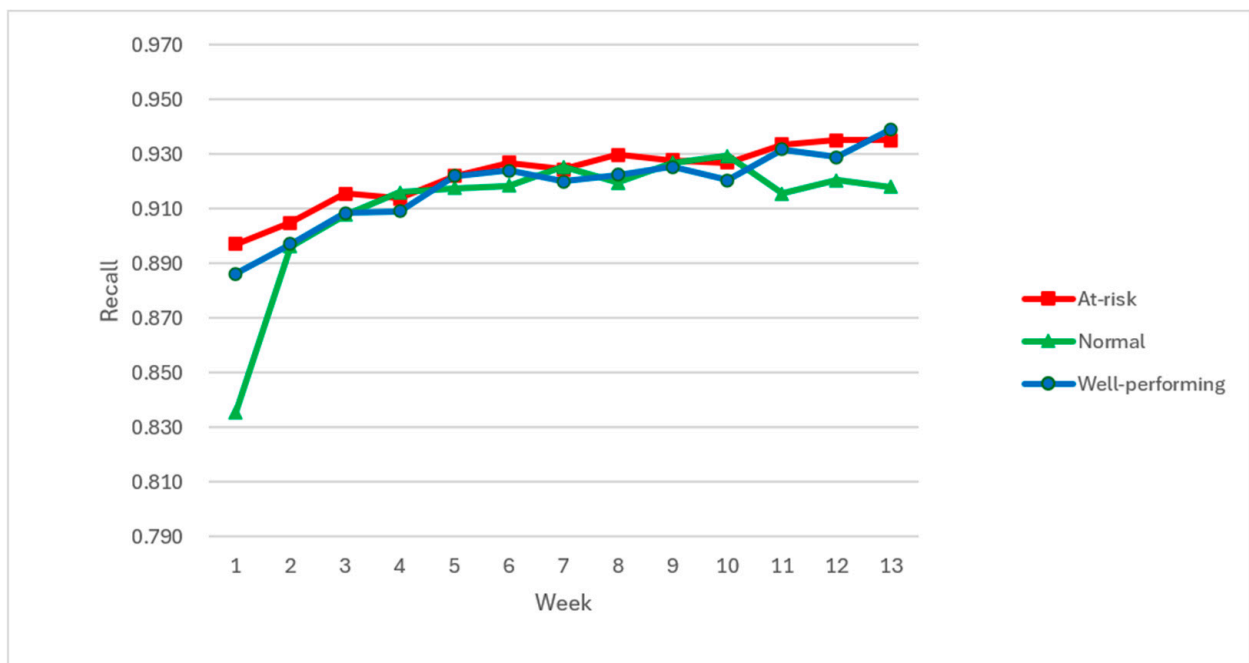


Figure 13. Per-class recall of classification for all 236 subjects.

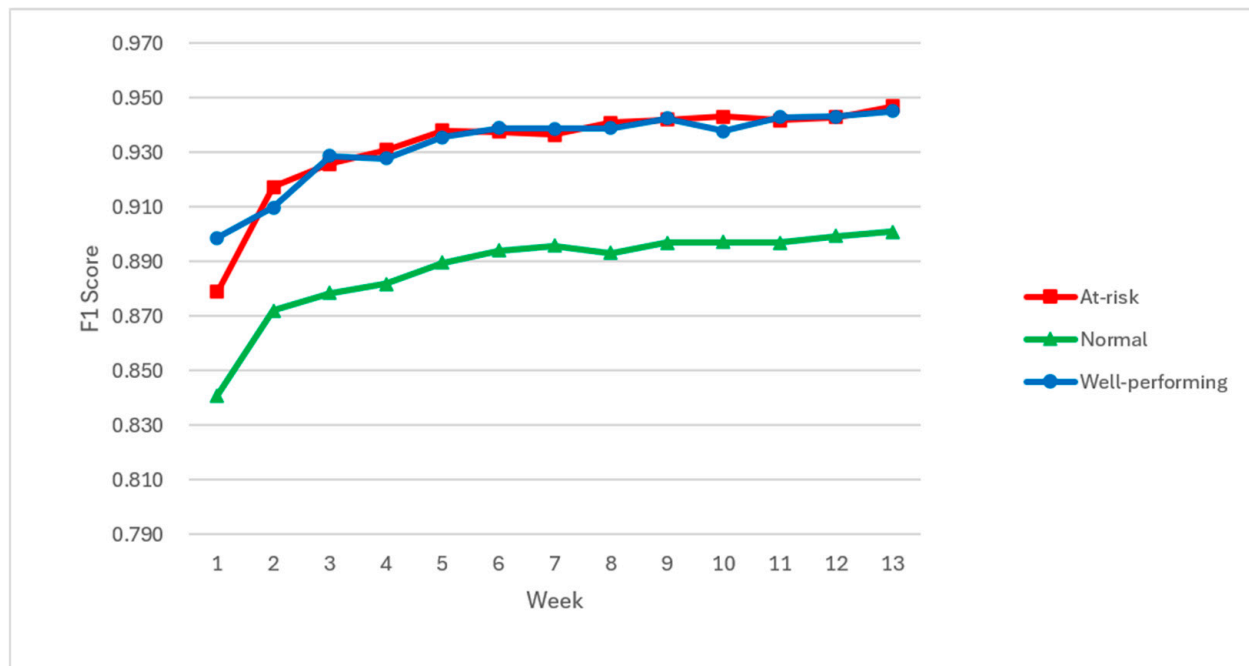


Figure 14. Per-class F1 score of classification for all 236 subjects.

5. Discussion

The results show that predictions using rank-based features can be as good as prior studies (Gray et al., 2016; Helal et al., 2018) that used common features such as raw or aggregated counts, transformed normalized values, or discretized categorical bins for prediction. Moreover, the compatibility of the rank-based features in classification is higher than in regression.

The classification-based prediction results from this study have important implications for educational practice. By employing sophisticated machine learning techniques, such as a Random Forest Classifier, the model demonstrated consistently strong prediction performance, achieving accuracy and F1 scores that consistently surpassed 90%. These metrics underscore not only the robustness of the model but also its practical utility in reliably identifying student outcome categories, including those students who may be at risk, as well as those on track for success. The high accuracy and F1 scores minimize the likelihood of misclassification, thereby supporting informed decision-making about potential early support strategies.

Table 6 presents a comparison of the features applied and the model performance in this research with similar studies. The evaluation metrics and results are as reported in each study. Due to differences in evaluation protocols, such as accuracy versus average accuracy and weekly versus aggregated reporting, the numerical results cannot be directly compared. Therefore, this comparison is intended to provide contextual positioning rather than a straightforward performance ranking. As summarized in Table 6, several prior approaches utilize richer feature sets or multiple data sources, with demographic data commonly included. In contrast, our method uses a more limited set of behavioral features while still achieving competitive performance.

The implementation of rank-based features enables the prediction model to adapt seamlessly to the diverse requirements and contexts of different academic disciplines. While rank-based transformations are conceptually simple, their systematic application as a cross-disciplinary standardization mechanism for early prediction across heterogeneous course designs has been largely overlooked in prior learning analytics research. Our

results demonstrate its underappreciated value in standardizing behavior data across diverse courses. This simplicity is a strength: it offers a transparent, easily adoptable method for institutions aiming to improve early predictions without sensitive data. This robust design ensures that educators, regardless of their teaching domain, can confidently apply the model within their own courses and expect consistently high performance and actionable insights.

Table 6. Comparison of the proposed approach with similar studies.

Research Studies	Features	Reported Performance	Scope
Adnan et al. (2021)	Demographics, Clickstream, Assessment	Random Forest Average accuracy = 59%, 79%, 84%, 88%, 90%, 91% at 0%, 20%, 40%, 60%, 80%, 100% of course length	Online study mode Not clearly specified: Discipline; With/without exam
Alturki et al. (2022)	Demographics, Post-enrollment features	Random Forest Accuracy: 77–94%	Single discipline; With exam; Not clearly specified: Study mode
Alwarthan et al. (2022)	Demographic data, Pre-University data, Courses grades, Cumulative GPA, Assessments marks	Random Forest Accuracy– Dataset 1: 99.662% Dataset 2: 91.715% Dataset 3: 95.363%	Single discipline; With exam; Not clearly specified: Study mode
Gonzalez-Nucamendi et al. (2023)	Demographic data, Pre-admission data, Admission data, Financing data	Random Forest Precision– Total sample: 92% Dropped out: 13.2% Retained: 99.4%	Not clearly specified: Discipline; Study mode; With/without exam
This Research	LMS log data (converted into rank-based features), Subject-specific features	Random Forest Accuracy: 87–93%	Across disciplines; Diverse study modes (Face-to-face, Online, and Blended); Both with and without exam

The predictive strength of these rank-based features is compellingly demonstrated through rigorous empirical evaluation. When employed within the subject performance classification framework, the model consistently achieves accuracy and F1 scores exceeding 80%. This robust performance underscores the discriminative power of the rank-based features. They are able to encapsulate patterns and distinctions that might elude more traditional metrics, resulting in a significant enhancement of the model’s ability to generalize across diverse student populations and academic scenarios. The model integrates three additional subject-specific features: EXAM, MODE, and DIVISION. These features complement the rank-based variables by capturing subject-dependent characteristics such as assessment structure, instructional mode, and disciplinary context. Their inclusion substantially improves model performance. Both accuracy and F1 scores increase to above 90%. This improvement shows that combining generalizable rank-based features with contextual information enhances predictive effectiveness.

A significant aspect of these findings is the ability of the predictive model to generate reliable forecasts as early as the third week of the semester. This early prediction window is critical in the context of academic support and student success strategies. Detecting potential academic challenges at such an incipient phase of the semester enables educators, advisors, and support staff to intervene proactively. The timeliness of early intervention can play a pivotal role in altering a student’s academic trajectory, potentially transforming outcomes that might otherwise have resulted in underperformance or course withdrawal.

Importantly, while this study did not implement interventions, the proposed model provides instructors with a scalable early warning foundation, enabling timely pedagogical actions such as targeted outreach, personalized feedback, or other proactive support to at-risk students.

The broader implications of these results are substantial and wide-reaching. On an institutional level, early identification and targeted support can lead to higher overall retention rates, improved graduation statistics, and enhanced reputational standing for the academic institution. More importantly, this predictive framework supports a shift towards a more inclusive and student-centered educational environment. Rather than relying solely on traditional, often reactive measures, educators are now empowered with data-driven insights that inform proactive interventions. This approach ensures that at-risk students are not left behind and that high-achieving students are encouraged to continue excelling.

It is essential to highlight that our predictive model relies exclusively on rank-based features and subject-specific variables, without incorporating any student demographic or sensitive personal data. By concentrating solely on learning behaviors and subject context, our design advances AI-driven educational analytics in a manner that is both ethical and responsible. Furthermore, the model utilizes data sources that are accessible to all instructors, as each can obtain student activity log files for their respective subjects from Blackboard. This facilitates rapid implementation and enables educators across various disciplines to leverage the predictive capabilities of the model effectively.

Our study also extends the learning analytics literature by showing that even a relatively limited set of LMS log variables can support strong prediction under the present learning context. This suggests that predictive performance depends not only on the richness of the logs but also on how the data are structured, processed, and aligned with the outcome of interest. Unlike much prior work (Conijn et al., 2017; López-Zambrano et al., 2020), which reports generalizability challenges across settings, our results indicate that high accuracy can be obtained under carefully specified modeling conditions when the behavioral signals in the LMS are closely related to the target outcome and when the modeling pipeline is carefully specified. These findings, therefore, contribute to ongoing debates in learning analytics about the practical value of digital trace data and the conditions under which such data can support reliable prediction.

6. Limitations and Future Research Directions

This study faced two main limitations, both of which were external to the research design and beyond the researchers' control. The first limitation relates to the format and type of data available from the Blackboard platform. Blackboard is a proprietary LMS, so the researchers did not have direct access to the data. The activity logs provided by Blackboard are limited in granularity. For example, it is not clear what actions students take after clicking on the teaching materials, such as whether they are thoroughly engaging with the material or merely opening files and leaving them idle in their browser for extended periods. Additionally, there is limited data regarding session durations and click timings, both of which are crucial for analyzing patterns such as click frequency at different stages of a session.

The second limitation concerns the data granularity of the marksheets provided by the institution. These marksheets adhere to a standardized format that captures only high-level assessment components, such as "test," "assignment," or "group project," without further detail. For example, an item labeled "test" might encompass multiple smaller tests conducted at different times, yet this distinction is not reflected in the data. Similarly, the marksheets do not specify whether assignments were conducted online under timed conditions, assigned as take-home projects, or divided into smaller components to track

ongoing learning progress. These omissions constrain a more detailed analysis of how assessment design impacts academic performance.

Future research should address these external limitations by advocating for enhanced data collection and reporting practices. In particular, Blackboard data collection should include more activity details to enhance the granularity of student activity analysis. Such comprehensive data collection will facilitate a deeper understanding of student engagement patterns and enable the development of more nuanced predictive models. Moreover, subsequent studies should seek to obtain richer assessment data, capturing the varied formats and structures of assignments. This refinement will support more accurate modeling of academic performance and foster insights into the relationship between assessment design and student outcomes.

Additionally, future research should focus on designing and evaluating tailored early intervention strategies prompted by predictive analytics. By systematically assessing the impact of personalized feedback, proactive academic advising, and adaptive learning resources, educational institutions can maximize the utility of early detection and advance student success. The integration of detailed data analysis with responsive interventions will further contribute to the creation of inclusive and supportive learning environments, thereby promoting the ongoing transformation and effectiveness of higher education.

7. Conclusions

This study presented a privacy-conscious and generalizable AI-based early-warning framework for identifying at-risk students using LMS engagement data only. The analysis is based on data from 236 subjects across four academic divisions. The results show that reliable early detection of academic risk is feasible as early as Week 3 of the semester. This finding holds even in highly heterogeneous teaching and learning contexts. A central contribution of this work lies in its emphasis on cross-disciplinary generalizability. The study also adopts a privacy-conscious design by using LMS engagement data only. The use of rank-based engagement features enables meaningful comparison of student behaviors across subjects with diverse instructional designs, assessment structures, and study modes, without relying on sensitive demographic or academic history data. While technically simple, this approach offers a transparent and scalable solution to a longstanding challenge in learning analytics research, namely, the limited transferability of early-prediction models beyond narrowly defined courses or disciplines.

Rather than claiming intervention effectiveness, the proposed model is positioned as a foundational early-warning mechanism that can support proactive and responsive teaching practices. Early identification of students who may be at risk provides instructors and academic support teams with actionable information. This information may enable timely pedagogical actions, such as targeted feedback, early advising, or referrals to learning support services. In this way, the model contributes to enhancing teaching responsiveness and student support capacity. This aligns with the broader goal of advancing teaching and learning through data-informed decision-making in higher education. The study also underscores the importance of ethical and responsible AI deployment in educational settings. The approach deliberately limits data inputs to engagement traces that are readily available in institutional LMS platforms. This design reduces privacy risks and mitigates potential bias associated with demographic profiling. It also lowers barriers to adoption across institutions with varying data infrastructures. While this study focused on early prediction, further research is needed to investigate how different pedagogical interventions triggered by early warnings influence student learning outcomes.

In conclusion, this research reframes early risk prediction as an enabling foundation for proactive, ethical, and scalable student support, rather than an end in itself. By prioritizing

generalizability, privacy, and pedagogical relevance, the study contributes practical insights to the growing body of research on AI-enhanced teaching and learning in higher education.

Author Contributions: Conceptualization, H.-S.C.; methodology, H.-S.C., A.W. and T.-L.W.; software, T.-L.W.; validation, H.-S.C.; formal analysis, H.-S.C. and T.-L.W.; investigation, H.-S.C.; resources, H.-S.C.; data curation, H.-S.C. and T.-L.W.; writing—original draft preparation, H.-S.C. and T.-L.W.; writing—review and editing, H.-S.C.; visualization, H.-S.C. and T.-L.W.; supervision, H.-S.C. and A.W.; project administration, H.-S.C.; funding acquisition, H.-S.C. All authors have read and agreed to the published version of the manuscript.

Funding: The work described in this paper was fully supported by a grant from the Research Grants Council of the Hong Kong Special Administrative Region, China (Project No., UGC/FDS24/H06/23).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this article are not readily available due to privacy and ethical restrictions.

Acknowledgments: During the preparation of this manuscript, the authors used Microsoft Copilot (version 1.25054.80.0) for the purposes of proofreading. The authors have reviewed and edited the output and take full responsibility for the content of this publication.

Conflicts of Interest: The authors declared no conflicts of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

References

- Adnan, M., Habib, A., Ashraf, J., Mussadiq, S., Raza, A. A., Abid, M., Bashir, M., & Khan, S. U. (2021). Predicting at-risk students at different percentages of course length for early intervention using machine learning models. *IEEE Access*, 9, 7519–7539. [\[CrossRef\]](#)
- Akçapınar, G., Altun, A., & Aşkar, P. (2019a). Using learning analytics to develop early-warning system for at-risk students. *International Journal of Educational Technology in Higher Education*, 16(1), 40. [\[CrossRef\]](#)
- Akçapınar, G., Hasnine, M. N., Majumdar, R., Flanagan, B., & Ogata, H. (2019b). Developing an early-warning system for spotting at-risk students by using eBook interaction logs. *Smart Learning Environments*, 6(1), 4. [\[CrossRef\]](#)
- Al-Shabandar, R., Hussain, A. J., Liatsis, P., & Keight, R. (2019). Detecting at-risk students with early interventions using machine learning techniques. *IEEE Access*, 7, 149464–149478. [\[CrossRef\]](#)
- Al-Sulami, A., Al-Masre, M., & Al-Malki, N. (2023). Predicting at-risk students' performance based on LMS activity using deep learning. *International Journal of Advanced Computer Science and Applications (IJACSA)*, 14(6), 1210–1220. [\[CrossRef\]](#)
- Alturki, S., Cohausz, L., & Stuckenschmidt, H. (2022). Predicting master's students' academic performance: An empirical study in Germany. *Smart Learning Environments*, 9(1), 38. [\[CrossRef\]](#) [\[PubMed\]](#)
- Alwarthan, S., Aslam, N., & Khan, I. U. (2022). An explainable model for identifying at-risk student at higher education. *IEEE Access*, 10, 107649–107668. [\[CrossRef\]](#)
- Alyahyan, E., & Dustegor, D. (2020). Predicting academic success in higher education: Literature review and best practices. *International Journal of Educational Technology in Higher Education*, 17, 3. [\[CrossRef\]](#)
- Bainbridge, J., Melitski, J., Zahradnik, A., Lauría, E. J. M., Jayaprakash, S., & Baron, J. (2015). Using learning analytics to predict at-risk students in online graduate public affairs and administration education. *Journal of Public Affairs Education*, 21(2), 247–262. [\[CrossRef\]](#)
- Bañeres, D., Rodríguez, M. E., Guerrero-Roldán, A. E., & Karadeniz, A. (2020). An early warning system to detect at-risk students in online higher education. *Applied Sciences*, 10(13), 4427. [\[CrossRef\]](#)
- Bañeres, D., Rodríguez-González, M. E., Guerrero-Roldán, A. E., & Cortadas, P. (2023). An early warning system to identify and intervene online dropout learners. *International Journal of Educational Technology in Higher Education*, 20(1), 3. [\[CrossRef\]](#)
- Bozkurt, A., & Sharma, R. C. (2020). Emergency remote teaching in a time of global crisis due to CoronaVirus pandemic. *Asian Journal of Distance Education*, 15(1), 1–6. [\[CrossRef\]](#)
- Bravo-Agapito, J., Romero, S. J., & Pamplona, S. (2021). Early prediction of undergraduate student's academic performance in completely online learning: A five-year study. *Computers in Human Behavior*, 115, 106595. [\[CrossRef\]](#)

- Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2002). SMOTE: Synthetic minority over-sampling technique. *Journal of Artificial Intelligence Research (JAIR)*, 16(1), 321–357. [\[CrossRef\]](#)
- Chiu, H. S., Ng, Z. T. B., & Cheung, Y. L. (2023). Academic performance prediction using artificial neural networks based on course-related student data. *International Journal of Advance Computational Engineering and Networking (IJACEN)*, 11(8), 15–20.
- Chiu, H. S., Wong, A., & Wong, T. L. (2025, July 10–14). *Data-driven analysis of learning behaviors among at-risk students across disciplines using data mining techniques*. The European Conference on Education (ECE2025), London, UK. [\[CrossRef\]](#)
- Chong, K. T., Ibrahim, N., Huspi, S. H., Kadir, W. M. N. W., & Isa, M. A. (2025). A systematic review of machine learning techniques for predicting student engagement in higher education online learning. *Journal of Information Technology Education: Research*, 24, 005. [\[CrossRef\]](#) [\[PubMed\]](#)
- Conijn, R., Snijders, C., Kleingeld, A., & Matzat, U. (2017). Predicting student performance from LMS data: A comparison of 17 blended courses using moodle LMS. *IEEE Transactions on Learning Technologies*, 10(1), 17–29. [\[CrossRef\]](#)
- Dai, W., Lin, J., Jin, F. J.-Y., Tsai, Y.-S., Srivastava, N., Le Bodic, P., Gašević, D., & Chen, G. (2025). Learning analytics for early identification of at-risk students and feedback intervention. *Journal of Learning Analytics*, 12(3), 102–125. [\[CrossRef\]](#)
- Daniel, S. J. (2020). Education and the COVID-19 pandemic. *Prospects*, 49, 91–96. [\[CrossRef\]](#) [\[PubMed\]](#)
- Dien, T. T., Luu, S. H., Thanh-Hai, N., & Thai-Nghe, N. (2020). Deep learning with data transformation and factor analysis for student performance prediction. *International Journal of Advanced Computer Science and Applications*, 11(8), 711–721. [\[CrossRef\]](#)
- Ding, M., Wang, Y., Hemberg, E., & O'Reilly, U.-M. (2019, March 4–8). *Transfer learning using representation learning in massive open online courses*. 9th International Conference on Learning Analytics & Knowledge (LAK19), Tempe, AZ, USA. [\[CrossRef\]](#)
- Fachola, C., Tornaría, A., Bermolen, P., Capdehourat, G., Etcheverry, L., & Fariello, M. I. (2023). Federated learning for data analytics in education. *Data*, 8(2), 43. [\[CrossRef\]](#)
- Foster, E., & Siddle, R. (2020). The effectiveness of learning analytics for identifying at-risk students in higher education. *Assessment & Evaluation in Higher Education*, 45(S1), 842–854. [\[CrossRef\]](#)
- Gonzalez-Nucamendi, A., Noguez, J., Neri, L., Robledo-Rella, V., & García-Castelán, R. M. G. (2023). Predictive analytics study to determine undergraduate students at risk of dropout. *Frontiers in Education*, 8, 1244686. [\[CrossRef\]](#)
- Gray, G., McGuinness, C., Owende, P., & Hofmann, M. (2016). Learning factor models of students at risk of failing in the early stage of tertiary education. *Journal of Learning Analytics*, 3(2), 330–372. [\[CrossRef\]](#)
- Hanna, M. G., Pantanowitz, L., Jackson, B., Palmer, O., Visweswaran, S., Pantanowitz, J., Deebajah, M., & Rashidi, H. H. (2025). Ethical and bias considerations in artificial intelligence/machine learning. *Modern Pathology*, 38(3), 100686. [\[CrossRef\]](#) [\[PubMed\]](#)
- Helal, S., Li, J., Liu, L., Ebrahimie, E., Dawson, S., Murray, D. J., & Long, Q. (2018). Predicting academic performance by considering student heterogeneity. *Knowledge-Based Systems*, 161, 134–146. [\[CrossRef\]](#)
- Hu, Y. H., Lo, C. L., & Shih, S. P. (2014). Developing early warning systems to predict students' online learning performance. *Computers in Human Behavior*, 36, 469–478. [\[CrossRef\]](#)
- Ifenthaler, D., & Yau, J. Y.-K. (2020). Utilising learning analytics to support study success in higher education: A systematic review. *Educational Technology Research and Development*, 68(4), 1961–1990. [\[CrossRef\]](#)
- Jamjoom, M., Alabdulkureem, E., Hadjouni, M., Karim, F., & Qarh, M. (2021). Early prediction for at-risk students in an introductory programming course based on student self-efficacy. *Informatica*, 45(6), 1–9. [\[CrossRef\]](#)
- Jang, Y., Choi, S., Jung, H., & Kim, H. (2022). Practical early prediction of students' performance using machine learning and eXplainable AI. *Education and Information Technologies*, 27(9), 12855–12889. [\[CrossRef\]](#)
- Karalar, H., Kapucu, C., & Gürüler, H. (2021). Predicting students at risk of academic failure using ensemble model during pandemic in a distance learning system. *International Journal of Educational Technology in Higher Education*, 18(1), 63. [\[CrossRef\]](#) [\[PubMed\]](#)
- Kumar, N., Chithrakumar, T., Thangarasan, T., Dhanasekar, J., & Logamurthy, P. (2025). AI-powered early detection and prevention system for student dropout risk. *International Journal of Computational and Experimental Science and Engineering (IJCESSEN)*, 11(1), 78–86. [\[CrossRef\]](#)
- Lakkaraju, H., Aguiar, E., Shan, C., Miller, D., Bhanpuri, N., Ghani, R., & Addison, K. L. (2015, August 10–13). *A machine learning framework to identify students at risk of adverse academic outcomes*. 21st ACM International Conference on Knowledge Discovery and Data Mining (pp. 1909–1918), Sydney, Australia. [\[CrossRef\]](#)
- Li, K. C., Wong, B., & Chan, H. T. (2023). Prediction of at-risk students using learning analytics: A literature review. In *Technology in education. Innovative practices for the new normal (ICTE 2023)*. Springer. [\[CrossRef\]](#) [\[PubMed\]](#)
- Liu, Q., Shakya, R., Khalil, M., & Jovanovic, J. (2025, March 3–7). *Advancing privacy in learning analytics using differential privacy*. 15th International Learning Analytics and Knowledge Conference (LAK25), Dublin, Ireland. [\[CrossRef\]](#)
- Liu, T., Wang, C., Chang, L., & Gu, T. (2022). Predicting high-risk students using learning behavior. *Mathematics*, 10(14), 2483. [\[CrossRef\]](#)
- López-Zambrano, J., Lara, J. A., & Romero, C. (2020). Towards portability of models for predicting students' final performance in university courses starting from Moodle logs. *Applied Sciences*, 10(1), 354. [\[CrossRef\]](#)
- Luo, X. (2024). Learning analytics based on big data: Student behavior prediction and personalized educational strategy formulation. *Applied and Computational Engineering*, 116, 7–13. [\[CrossRef\]](#)

- Means, B., & Neisler, J. (2021). Teaching and learning in the time of COVID: The student perspective. *Online Learning*, 25(1), 8–27. [\[CrossRef\]](#)
- Murphy, M. P. A. (2020). COVID-19 and emergency eLearning: Consequences of the securitization of higher education for post-pandemic pedagogy. *Contemporary Security Policy*, 41(3), 492–505. [\[CrossRef\]](#)
- Nimy, E., Mosia, M., & Chibaya, C. (2023). Identifying at-risk students for early intervention—A probabilistic machine learning approach. *Applied Sciences*, 13(6), 3869. [\[CrossRef\]](#)
- Russell, J.-E., Smith, A., Larsen, R. G., & Larsen, R. G. (2020). Elements of success: Supporting at-risk student resilience through learning analytics. *Computers & Education*, 152, 103890. [\[CrossRef\]](#)
- Sneyers, E., & De Witte, K. (2018). Interventions in higher education and their effect on student success: A meta-analysis. *Educational Review*, 70(2), 208–228. [\[CrossRef\]](#)
- Thammasiri, D., Delen, D., Meesad, P., & Kasap, N. (2014). A critical assessment of imbalanced class distribution problem: The case of predicting freshmen student attrition. *Expert Systems with Applications*, 41(2), 321–330. [\[CrossRef\]](#)
- Topali, P., Ortega-Arranz, A., Rodríguez-Triana, M. J., Er, E., Khalil, M., & Akçapınar, G. (2025). Designing human-centered learning analytics and artificial intelligence in education solutions: A systematic literature review. *Behaviour & Information Technology*, 44(5), 1071–1098. [\[CrossRef\]](#)
- Tsiakmaki, M., Kostopoulos, G., Kotsiantis, S., & Ragos, O. (2020). Transfer learning from deep neural networks for predicting student performance. *Applied Science*, 10, 2145. [\[CrossRef\]](#)
- Ujkani, B., Minkovska, D., & Hinov, N. (2024). Course success prediction and early identification of at-risk students using explainable artificial intelligence. *Electronics*, 13(21), 4157. [\[CrossRef\]](#)
- Vaarma, M., & Li, H. (2024). Predicting student dropouts with machine learning: An empirical study in Finnish higher education. *Technology in Society*, 76, 102474. [\[CrossRef\]](#)
- Waheed, H., Hassan, S., Aljohani, N. R., Hardman, J., Alelyani, S., & Nawaz, R. (2020). Predicting academic performance of students from VLE big data using deep learning models. *Computers in Human Behavior*, 104, 106189. [\[CrossRef\]](#)
- Wang, S., Wang, F., Zhu, Z., Wang, J., Tran, T., & Du, Z. (2024). Artificial intelligence in education: A systematic literature review. *Expert Systems with Applications*, 252(A), 124167. [\[CrossRef\]](#)
- Wong, A., So, J., & Ng, Z. T. B. (2024). Developing a web application for Chinese calligraphy learners using convolutional neural network and scale invariant feature transform. *Computers and Education: Artificial Intelligence*, 6, 100200. [\[CrossRef\]](#)
- Wong, A., So, J., Tsang, K., & Wei, R. (2023). Selecting the best K features for predicting student participation in generic competency development activities in higher education. *International Journal of Emerging Technologies in Learning (ijET)*, 18(23), 197–213. [\[CrossRef\]](#)
- Wong, B. T. M., & Li, K. C. (2020). A review of learning analytics intervention in higher education (2011–2018). *Journal of Computers in Education*, 7(1), 7–28. [\[CrossRef\]](#)
- Zhao, Y., Wang, J., Tan, X., Wen, L., Gao, Q., & Wang, W. (2025). Privacy-preserving and interpretable grade prediction: A differential privacy integrated TabNet framework. *Electronics*, 14(12), 2328. [\[CrossRef\]](#)

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.