JOURNAL OF LEARNING ANALYTICS

Volume 12(3), 102–125. https://doi.org/10.18608/jla.2025.8735

Learning Analytics for Early Identification of At-Risk Students and Feedback Intervention

Wei Dai[1], Jionghao Lin[2], Flora Ji-Yoon Jin [3], Yi-Shan Tsai[4], Namrata Srivastava[5], Pierre Le Bodic[6], Dragan Gašević[7], and Guanliang Chen[8]*

Abstract

Supporting academically at-risk students has attracted much attention in the field of learning analytics. However, much of the research in this area has focused on developing advanced machine learning models to predict students' academic performance, which alone is insufficient to improve student learning without the implementation of timely interventions. Among the studies that attempted to mitigate this limitation by deploying intervention feedback to enhance learning, few created their feedback based on established theories of effective feedback. This theoretical oversight may limit students' uptake of the provided intervention. In response to these gaps, we conducted a study that aimed at supporting at-risk students at the early stage of an undergraduate-level course. Specifically, we developed predictive machine learning models using trace and academic data from the previous offering of a course and applied these models to identify at-risk students in the subsequent semester's offering of the same course. For the identified at-risk students, we sent intervention emails designed by feedback experts based on a relational feedback framework designed to enhance feedback effectiveness by strengthening student-instructor relationships. We evaluated the effectiveness of the proposed approach by assessing the performance of the predictive models in terms of generalizability and measuring the impact of the feedback intervention on students' behavioural engagement in learning. Results showed that (i) our predictive models demonstrated a high prediction accuracy (with AUC scores above 0.8) when applied to a new cohort of students; (ii) more than 30% of the identified at-risk students visited previously unengaged learning activities within two weeks following the intervention; and (iii) survey responses from 9.27% of at-risk students indicated general satisfaction with the provided feedback intervention, and 60% of the respondents expressed a preference for receiving the intervention more frequently than the twice-per-semester frequency implemented in the present study.

Notes for Practice

- Integrating predictive models with relational feedback interventions into learning platforms is encouraged to help detect and support academically at-risk students, thereby enhancing their behavioural engagement with learning activities at the early stages of a course.
- The accuracy of predictive models for identifying at-risk students often declines when applied to new student cohorts without retraining. To maintain reliable performance across diverse student populations, these models should be regularly updated using data from new cohorts.
- Predictively identified at-risk students responded positively to the relational feedback interventions. However, as our survey results indicate, preferences for intervention frequency vary. Therefore, students should be given flexible options regarding whether and when to receive this support.

Keywords

Learning analytics, predictive modelling, model generalization ability, early feedback intervention, effective feedback design.

Submitted: 22/10/2024 — **Accepted:** 23/09/2025 — **Published:** 30/11/2025

[1] *Email: wei.dai1@monash.edu, daiw@hku.hk Address: Faculty of Information Technology, Monash University, Melbourne, Australia. ORCID iD: https://orcid.org/0000-0001-7206-9347*

[2] *Email: jionghao.lin@monash.edu, jionghao@hku.hk Address: Faculty of Information Technology, Monash University, Melbourne, Australia. Faculty of Education, University of Hong Kong, Hong Kong SAR, China. ORCID iD: https://orcid.org/0000-0003-3320-3907*

[3] *Email: flora.jin@monash.edu Address: Faculty of Information Technology, Monash University, Melbourne, Australia. ORCID iD: https://orcid.org/0009-0002-1825-4028*

[4] *Email: yi-shan.tsai@monash.edu Address: Faculty of Information Technology, Monash University, Melbourne, Australia. ORCID iD: https://orcid.org/0000-0001-8967-5327*

[5] *Email: namrata.srivastava@vanderbilt.edu Address: Vanderbilt University, Nashville, Tennessee, United States. ORCID iD:*

https://orcid.org/0000-0003-4194-318X
[6] Email: pierre.lebodic@monash.edu Address: Faculty of Information Technology, Monash University, Melbourne, Australia. ORCID iD: https://orcid.org/0000-0003-0842-9533
[7] Email: dragan.gasevic@monash.edu Address: Faculty of Information Technology, Monash University, Melbourne, Australia. ORCID iD: https://orcid.org/0000-0001-9265-1908
[8]*Corresponding author. Email: guanliang.chen@monash.edu Address: Faculty of Information Technology, Monash University, Melbourne, Australia. ORCID iD: https://orcid.org/0000-0002-8236-3133

1. Introduction

Supporting students at risk of failing courses has gained much attention in higher education due to the multifaceted impacts of their non-completion (Veerasamy et al., 2020). Foundational courses, such as the introductory programming course examined in this study, are pivotal as they equip students with skills that are essential for advanced subjects in subsequent academic years. Failure to master either foundational or advanced courses threatens students' ability to successfully graduate, leading to potential dropout. Student dropout can have significant consequences not only for the individual student but also for the university and society at large. At the student level, withdrawal from college or university may lead to feelings of failure and diminished self-confidence, which could adversely affect students' mental health (Faas et al., 2018; Berka & Marek, 2021). In addition, the lifetime earnings of individuals without a degree are significantly lower than those of college/university graduates (Singell & Waddell, 2010; Adnan et al., 2021). At the university level, increased rates of student dropout may negatively impact institutional funding and complicate the allocation of educational resources, such as instructors and classrooms, for students who might not continue their studies in the upcoming semester (Berka & Marek, 2021). At the society level, higher-education graduates are more likely to contribute to tax revenues, productivity, and innovation, all of which are pivotal to the economic growth of a society (Gupta et al., 2020). Conversely, people lacking education are more prone to unemployment or underemployment, often necessitating reliance on government assistance. This dependence can escalate public expenditures on welfare and healthcare (Latif et al., 2015). Beyond economic impacts, higher dropout rates are indicative of lower national literacy levels, which correlate with increased crime rates and diminished civic participation, thereby undermining social stability and cohesion (Gupta et al., 2020). To mitigate the negative impacts associated with rising student dropout rates, it is essential to identify at-risk students early and implement timely interventions that encourage their continued engagement in their studies. However, using traditional manual approaches to identify these students has been a challenging task due to the increasing numbers of students and relatively stagnant resources available in higher-education institutions (Khalil et al., 2024).

The use of learning analytics has been widely recognized as a promising approach to addressing challenges of early identification of at-risk students in higher education (Campbell et al., 2007). Due to the widespread integration of online learning platforms (e.g., Moodle and Ed) into higher-education institutions, a massive volume of student learning data can be tracked, ranging from web page browsing to assignment submissions. The availability of such diverse data means it may be used in learning analytics to train predictive models and develop early warning systems that aim to support at-risk students (Zacharis, 2015; Akçapınar et al., 2019). Much of the research in learning analytics has focused on developing models that use features extracted from students' learning trace data to predict academic success at the end of a course (Bainbridge et al., 2015; Zacharis, 2015). While these predictions offer valuable insights into the factors influencing student success or failure, their practical utility is limited due to the timing of such predictive analysis (Waheed et al., 2023). If trained on the data collected at the end of the course, such predictive analysis may be unsuitable to inform interventions in the early stage of the course (Waheed et al., 2023; Brooks et al., 2015), thus failing to enhance student learning throughout the course. To support at-risk students, predictive modelling should be applied early in the semester (Brooks et al., 2015), allowing for the timely delivery of interventions. The earlier an intervention is deployed, the sooner a student will have the opportunity to adjust their learning approach to improve their academic performance (Y. Zhang et al., 2014).

An increasing number of studies aim to build predictive models that can be applied in the early stage of courses (Y. Zhang et al., 2014; Brooks et al., 2015; L. Zhang & Rangwala, 2018; Lu et al., 2017; Waheed et al., 2023). While accurate predictive models are crucial for detecting at-risk students, they are insufficient alone to increase course completion without being combined with effective intervention strategies (Clow, 2012). Some studies have attempted to offer interventions such as arranging in-person meetings with instructors to discuss learning progress (Lu et al., 2017), holding workshops for students to attend voluntarily (Winston et al., 2014), or making phone calls to students identified to be at risk (Dawson et al., 2017). These strategies, however, require significant time and effort from instructors and higher-education institutions, which may not always lead to large improvements (Dawson et al., 2017) and may not be scalable, especially in an era of expanding numbers of higher-education students. A relatively simple and cost-effective approach is sending short messages via email as feedback to inform students of their learning progress (Wong & Li, 2020). The inherent scalability of email allows instructors to communicate with a growing number of students without additional costs and enables integration with learning management systems (LMSs) for automated alerts based on specific triggers such as declining engagement with learning materials. Sending email messages has been demonstrated to be effective in encouraging students to seek help and change their learning behaviour by increasing the interactions between students and instructors (Jayaprakash et al., 2014). The success of email feedback depends largely on the design of the feedback, including the content it encompasses and the manner in which this content is delivered (Hattie & Timperley, 2007). For example, according to the design framework of relational feedback (Dai et al., 2025), feedback that contains actionable information and is phrased politely is beneficial for enhancing student-instructor relationships, which may contribute to students' utilization of the feedback. However, little attention has been paid to the design and effectiveness of feedback interventions based on prominent feedback theories for an improved uptake of interventions.

In response, we conducted a study, detailed in this paper, aimed at supporting at-risk students in an undergraduate-level course on algorithms and programming which was delivered in the first semester of 2023. In this study, "at-risk students" are defined as those predicted to receive a final course grade below 50% (i.e., "fail"), which is the institutional threshold for a failing grade. Specifically, we developed predictive models based on student learning data recorded by online learning platforms during a historical delivery of the course (i.e., the first semester of 2022). These models were then applied to identify at-risk students in a new semester (i.e., the first semester of 2023). For those predicted at-risk students, we sent emails with feedback on the students' online learning participation to encourage their behavioural engagement in learning activities. These feedback emails were crafted based on the design framework of relational feedback (Dai et al., 2025) by researchers with expertise in feedback design. We evaluated the effectiveness of our approach, i.e., early prediction of at-risk students combined with a feedback intervention, by addressing the following research questions:

RQ1: To what extent can predictive models trained on one semester accurately generalize to identify at-risk students in a subsequent semester of the same course?

RQ2: To what extent, if any, does students' engagement in the online learning platform change after receiving the feedback intervention?

RQ3: How do at-risk students perceive the provided feedback intervention on their online learning engagement?

For RQ1, we evaluated the performance of the predictive model in terms of generalizability, i.e., whether a model constructed with data collected from a semester can be applied to identify at-risk students in another semester, an area that remains relatively underexplored. Predictive models trained on data from one semester could exhibit diminished performance when applied to a subsequent semester for identifying at-risk students. This degradation can be attributed to covariance drift and concept drift (Sonnleitner et al., 2025). Covariance drift denotes changes in the distribution of the input features, such as student demographics or learning-engagement indicators, between the training and test datasets. Concept drift, in contrast, refers to the shifts in the underlying relationship between the input features (e.g., students' learning engagement indicators and their demographics) and the target outcome (e.g., whether students are at risk or not). These drifts are practically inevitable in educational settings, as institutions may modify curricula, student cohorts differ year to year, and external events (e.g., a pandemic) can abruptly reshape learning behaviours and success factors. In this context, by generalizability we refer to a model's ability to maintain acceptable predictive accuracy across cohorts despite these inevitable drifts, without retraining. Investigating this kind of generalizability is essential for early warning systems, particularly in real-world educational contexts where comprehensive historical data for model retraining are often limited (López-Zambrano et al., 2020). A model with such generalizability eliminates the need for extensive retraining, thereby facilitating the timely identification of and intervention for at-risk students. To answer RQ1, we evaluated the performance of predictive models using datasets derived from two different semesters' offerings of the same course. We also examined the differences in important features across the models trained on these diverse datasets.

To answer RQ2, we measured changes in student engagement with learning activities on the online learning platform following the analytics-informed feedback intervention. This was conducted to evaluate the effectiveness of such intervention strategies in fostering improved learning engagement.

To answer RQ3, we distributed a survey to at-risk students. Investigating student perceptions of analytics-informed feedback interventions (RQ3) may assist in refining feedback interventions for future students to achieve higher participation and engagement in learning. Student perceptions of analytics-informed feedback interventions are critical for understanding the extent to which interventions may motivate students and promote their increased engagement.

2. Literature Review

2.1 Early Identification of At-Risk Students

Identifying at-risk students at the early stage of a course is crucial for implementing timely interventions that can increase the completion rate in higher education. To address this challenge, numerous studies have been conducted to investigate the association between student characteristics and academic success to suggest indicators for early detection of at-risk students

(Bainbridge et al., 2015; Brooks et al., 2015; Chipchase et al., 2017; Waheed et al., 2023). The student characteristics examined in the literature mainly include (i) student demographics such as gender (Bainbridge et al., 2015; Na & Tasir, 2017; Nimy et al., 2023) and age (Bainbridge et al., 2015) and (ii) student learning behaviours such as the frequency of access to learning materials (Bainbridge et al., 2015; R. S. Baker et al., 2015; Waheed et al., 2023), the submission time of assignments (Chipchase et al., 2017; Falkner & Falkner, 2012; Osborne & Lang, 2023; Khan et al., 2023), and assignment performance (Bainbridge et al., 2015; R. S. Baker et al., 2015; Khan et al., 2023). For instance, Bainbridge and colleagues (2015) reported that students who visited course content less frequently achieved poorer performance in the final examination. Among various predictors related to student success, those derived from student engagement are the most significant contributors to prediction performance (Liz Domínguez et al., 2019). A meta-analysis conducted by Wang and Mousavi (2023) on the effects of different engagement indicators on academic achievement found that students who actively participate in learning activities are more likely to achieve higher grades.

In addition to identifying risk factors (e.g., the low engagement with course content (Bainbridge et al., 2015)) that contribute to student failure, some researchers focused on investigating the prediction performance of different models for identifying at-risk students. The majority of these studies framed the identification of at-risk students as a classification problem, aiming to distinguish between students likely to pass the course and those at risk of failing. Commonly used classification algorithms include logistic regression (LR) (L. Zhang & Rangwala, 2018; Masabo et al., 2023), tree-based models such as decision trees (Figueroa-Cañas & Sancho-Vinuesa, 2019; Masabo et al., 2023; López-García et al., 2023), random forest (RF) (Na & Tasir, 2017; Adnan et al., 2021; Masabo et al., 2023), and neural networks (Waheed et al., 2023; Christou et al., 2023). For example, Figueroa-Cañas and Sancho-Vinuesa (2019) adopted decision trees to detect dropout students at the deadline of each of the four continuous assessment tests due to their simplicity and interpretability. The results indicated that the model developed using data from the first day of the semester up to the third assessment submission deadline (before the halfway point of the semester) achieved optimal prediction performance with an accuracy of 90.2%. A comparative study (He et al., 2015) showed that LR outperformed decision tree and RF in predicting failure weekly, as measured by the area under the ROC curve (AUC). Beyond traditional classification models, a few studies aimed to develop novel models to enhance the accuracy of early prediction of student success. Er (2012) and Hu and colleagues (2014) employed different decision schemes to combine results from multiple classifiers. Experiments in both studies indicated that the ensembles of classifiers outperformed individual algorithms.

The aforementioned studies demonstrated the feasibility of predicting at-risk students throughout the duration of a course. However, the generalizability of these predictive models across different semesters of a course has not been thoroughly investigated (Brooks et al., 2015; Akçapınar et al., 2019). Such investigation is particularly important because the historical data required for model training may not always be available for new cohorts of students each semester. In addition, the instructional design of a course can vary significantly across different semesters. Differences in instructional conditions can have a significant impact on the accuracy of predicting at-risk students (Gašević et al., 2016). López-Zambrano and colleagues (2020) conducted a comprehensive analysis of predictive models' generalizability on 24 university courses. They created a predictive model for each of the 24 courses and applied each model to the other courses for testing. The results showed that most of the AUC values of the models in the new courses dramatically decreased, with some dropping to levels below randomness. Gardner and colleagues (2023) reported inconsistent performance of predictive models trained on a dataset from a source institution when applied to four additional datasets from other institutions. While models showed good performance for two of these institutions, they did not perform well for the other two. These mixed results highlight challenges in the generalizability of predictive models across different educational contexts. Therefore, in the present study, we aim to investigate whether the models developed by using data collected from a historical semester of a course can be effectively transferred or applied to predicting students at risk in a new semester of the same course.

2.2 Intervention for At-Risk Students

While accurately identifying at-risk students in the early days of a semester is crucial, providing effective interventions supporting their learning is even more critical. However, most of the existing research paid attention to the development of advanced predictive models and much less attention to the design and impact of interventions (Arnold & Pistilli, 2012; Dawson et al., 2017; López-García et al., 2023; Schmidt et al., 2025). Common interventions offered to students include notification messages sent by instructors (Lu et al., 2017), phone calls to students (Dawson et al., 2017), or early warning systems (Arnold & Pistilli, 2012; Foster & Siddle, 2020; Bañeres et al., 2020), which primarily inform students of their learning progress or performance. Apart from simple notifications, Jayaprakash and colleagues (2014) encouraged students to join an online learning support service that offered students access to additional learning materials and opportunities to meet with peer mentors and instructors. Similar strategies, i.e., instructors having face-to-face meetings with students who struggled with their studies, were employed by Lu and colleagues (2017) and Zhang and colleagues (2014). In the context of medical education, workshops were arranged for at-risk students to enhance their learning (Winston et al., 2014).

Although in-person meetings or workshops provide a platform for students to address their specific learning needs, these interventions can be highly time-consuming for educators, especially in the current landscape of higher education where student numbers are continuously increasing. Simple warnings about academic risks sent via messages have been shown to motivate students to seek help (Jayaprakash et al., 2014). These messages could be further enhanced by including more effective components such as actionable information that students can use to achieve their learning goals. However, existing research indicates that the use of analytics-based early warning systems does not result in the use of principles recommended by research and theories of feedback (Bañeres et al., 2020). Specifically, there is a notable lack of focus in the literature on the design of these intervention messages to maximize their effectiveness, which motivated the present study to examine the effectiveness of intervention messages based on research and theories of feedback (Borrella et al., 2022).

The effectiveness of intervention messages provided to at-risk students has been investigated in a number of studies (R. Baker et al., 2016; Borrella et al., 2019, 2022). The findings in these studies are consistent, which is that the email intervention had no statistically significant effect on reducing student dropout. For example, Borrella and colleagues (2022) divided at-risk students into a treatment group and a control group. Only the treatment group received the intervention, which was aimed at encouraging students to take the midterm exam. However, there was no significant difference in the percentages of students who took the midterm exam between the treatment and control groups. However, they identified at-risk students and provided interventions only once during a course, which might impact the effectiveness of the intervention provided. If the identification and intervention occur too early, they may not capture students who become at risk later. Conversely, if they happen too late, there might not be enough time for the intervention to take effect (Wong & Li, 2020). In addition, these studies focused on assessing the impact of email-based interventions on overall student retention or dropout rates but did not explore their effects on individual students. This omission highlights a gap in understanding how such interventions might influence specific learning behaviours at the individual level. To address these limitations, our study implemented the identification and intervention at two distinct points during a course. Regarding the investigation into the effectiveness of intervention, we specifically focused on the impact of the intervention on individual student engagement levels by comparing their engagement with learning activities before and after intervention.

2.3 Theoretical Frameworks for Designing Effective Feedback

Message intervention, as a type of feedback, can be strategically designed to enhance its effectiveness in promoting student learning. Numerous feedback theories and frameworks have been proposed to guide the creation of effective feedback (Hattie & Timperley, 2007; Ryan et al., 2021; Dai et al., 2025). For example, Hattie and Timperley (2007) suggested that feedback should contain information at task level (focusing on the performance of the task), process level (focusing on the strategies used to complete the task), self-regulation level (focusing on encouraging students to reflect on their learning), and self level (focusing on evaluating the student rather than the task). Unlike the four-level feedback framework proposed by Hattie and Timperley (2007), which positions instructors as managers in the feedback process, the learner-centred feedback framework emphasizes the students' active role in monitoring their learning (Ryan et al., 2021). To facilitate text-based learner-centred feedback, Ryan and colleagues (2021) identified feedback components for instructors' practical use, such as helping students obtain frequent feedback from various sources and promoting student independence.

Recently, there has been increasing research interest in relational feedback, which focuses on developing relationships between students and the learning community (e.g., peers, instructors, and institutions) through feedback (Kastberg et al., 2020; Dai et al., 2025). Establishing a rapport with peers and instructors is crucial for all students, as it enhances their learning experience by fostering a sense of community and support. Such positive feelings can encourage students to engage in the learning process (Middleton et al., 2023; Heron et al., 2023). However, building these relationships can be particularly challenging for at-risk students, who tend not to participate in learning activities (Dix et al., 2020). Their lower participation can lead to a distance between them and their instructors or peers, which might result in these students not paying attention to, or even disregarding, the feedback or interventions provided (Middleton et al., 2023). Neglecting feedback can further aggravate their academic difficulties (Heron et al., 2023). To develop student-instructor relationships and encourage at-risk students' engagement in learning, we adopted the framework for creating relational feedback proposed by Dai and colleagues (2025) to craft our feedback intervention. This framework suggests that relational feedback should (i) clarify students' performance, (ii) contain actionable information for students to improve their learning, (iii) encourage students to communicate with the learning community, and (iv) evoke students' positive emotions. The implementation of these relational principles in our feedback intervention is detailed in Section 3.3.

3. Methods

The study focused on identifying and supporting at-risk students within an undergraduate-level course teaching introductory programming at an Australian university. The timeline and workflow of our study are depicted in Figure 1. The study had two stages, i.e., training and implementation. In the training stage, we developed predictive models using data from students enrolled in the first semester of 2022 (referred to as the source semester). In the implementation stage, we applied predictive models in the first semester of 2023 (the target semester) to identify and subsequently support at-risk students through feedback

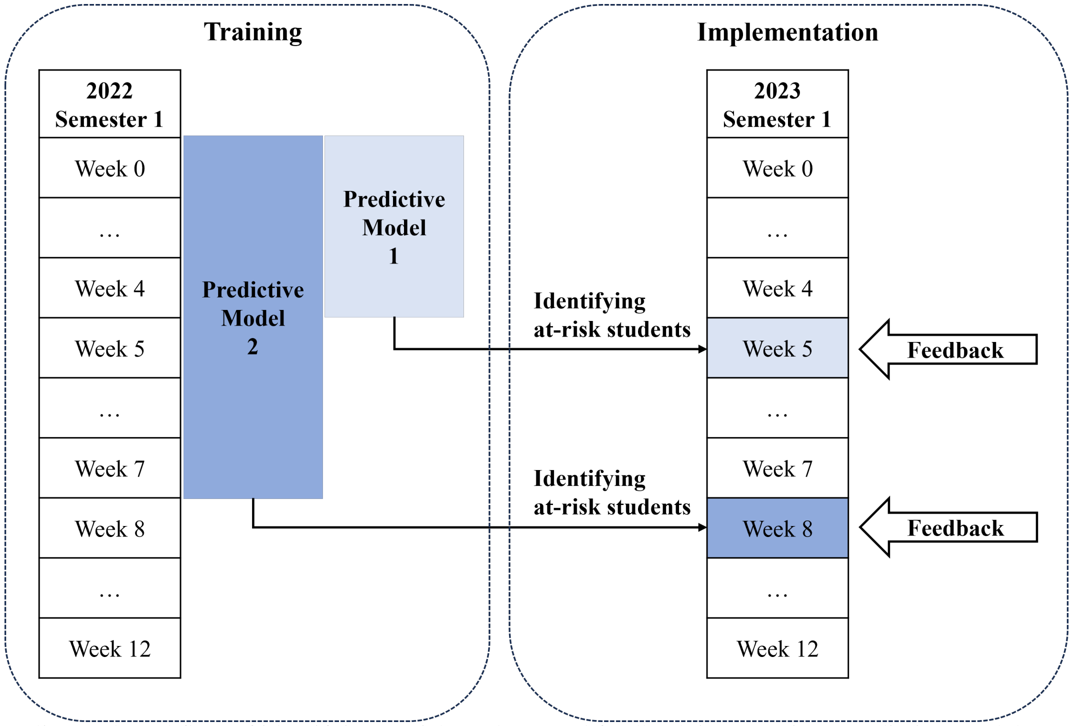


Figure 1. Timeline and workflow of identifying at-risk students and feedback intervention. Predictive Model 1 was trained on the data from week 0 to week 4 in the first semester of 2022 and was applied to identify at-risk students in week 5 in the first semester of 2023. Predictive Model 2 was trained on the data from week 0 to week 7 in the first semester of 2022 and was applied in week 8 in the first semester of 2023. The feedback intervention was implemented right after the identification of at-risk students.

interventions. In both semesters, educators used the Ed platform[1], an LMS that enables students to access learning materials, complete learning activities, and communicate with peers and instructors. Additionally, the Moodle LMS was used to manage assignment submissions and access grade reports. This course, consistent across both semesters, spanned 13 weeks from week 0 to week 12. Students were required to engage in four weekly tasks: (i) completing pre-class activities on Ed, (ii) attending the face-to-face workshop and completing workshop activities on Ed, (iii) participating in the face-to-face applied class and completing applied class activities on Ed, and (iv) completing post-class activities on Ed. The forthcoming subsections elucidate the details of our study, including developing predictive models, identifying at-risk students, and designing and delivering feedback interventions. Finally, we outline the methods used to answer the proposed research questions.

Our study obtained ethics approval from Monash University under project number 31325. Students' data from the Ed platform and Moodle LMS were retrieved through the university database under the university's consent framework, where students had previously been informed and had agreed that their de-identified learning data could be used for research. Because the feedback intervention served as direct pedagogical support, it remained within the bounds of this existing consent. For the survey delivered to ask students' perceptions on the feedback intervention, participants first reviewed an information sheet and then provided explicit electronic consent before responding. All data were de-identified and processed in accordance with the university's strict privacy safeguards throughout the project.

3.1 Predictive Model Construction

3.1.1 Data Collection and Feature Engineering

The data for training predictive models were collected from the Ed platform and Moodle LMS used by students who enrolled in the first semester of 2022. As the predictive models were applied to the first semester of 2023, using the data from the same semester of the previous year was deemed a suitable way to potentially minimize the impact of seasonal variations (e.g., public holidays) on the model's accuracy.

We developed the features for our predictive models by referencing the framework established by Jovanović and colleagues (2021), which defines a comprehensive set of indicators for measuring student engagement levels based on an extensive literature review. Features used in our predictive models are listed in Table 1. Four groups of features were extracted: (i)

[1] https://edstem.org/

general indicators of students' overall engagement with online learning platforms ("General indicators"), (ii) indicators of students' engagement tailored to different online learning actions ("Action-specific indicators"), (iii) general regularity of engagement indicators ("Regularity of general indicators"), and (iv) regularity of engagement indicators tailored to different online learning actions ("Regularity of action-specific indicators"). Specifically, there were six distinct online learning actions: (i) "forum contribution," denoting student interactions with the discussion forum on the Ed platform, such as posting or deleting messages; (ii) "forum consumption," denoting activities related to browsing or reading discussion forum content; (iii) "grade report," denoting the action of checking course grades; (iv) "course main page," denoting visits to the course main page for accessing course information, links, and updates; (v) "learning materials," denoting students' actions related to learning materials such as viewing lesson slides; and (vi) "assignment," denoting actions related to assignments, including reading assignment requirements and submitting assignments. Therefore, within the categories "Action-specific indicators" and "Regularity of action-specific indicators," there were two features for each of the six learning actions, resulting in a total of 12 features per category. In addition to these learning behavioural indicators, we enriched our models by incorporating demographic features of students, which include gender (e.g., male, female), age, years since degree commencement (e.g., 1 year, 2 years), degree type (e.g., single or double degree), citizenship status (e.g., domestic or international), and the number of different levels of grades (i.e., HD, D, C, P, and F) obtained in previous courses. The demographic information of students enrolled in the source and target semesters is detailed in Table 2. Since this study was conducted within a course offered by the Faculty of Information Technology and included all enrolled students, the observed gender imbalance reflects the actual enrolment demographics of the course and is consistent with broader gender distributions in STEM education.

3.1.2 Model Evaluation and Selection

As discussed in Section 2.1, LR (L. Zhang & Rangwala, 2018), RF (Na & Tasir, 2017), and gradient boosting tree (GBT) (Lin et al., 2023) have been widely used in the literature due to their effectiveness in the task of performance prediction. We used these three classifiers and evaluated their performance to select the optimal model for predicting at-risk students. LR is a statistical model commonly used for binary classification tasks, such as predicting student performance as "pass" or "fail" (Masabo et al., 2023). RF is an ensemble learning approach that outputs the prediction result by combining the decisions from multiple decision trees. GBT is a machine learning technique that builds tree models sequentially. Each new tree can correct errors made by previously trained trees.

To determine the optimal model for identifying at-risk students, we compared the performance of models—i.e., LR, RF, and GBT—weekly from week 0 to week 7 (the midterm of the source semester, i.e., the first semester of 2022), using data available up to the end of each respective week. For each week from 0 to 7, we randomly partitioned the data into an 80% training set and a 20% test set. The three models (i.e., LR, RF, and GBT) were trained on the training set and then evaluated on the test set using the AUC score. We computed the average AUC score for each model from week 0 to week 7 and selected the model with the highest average AUC score for performing the predictions in the target semester (i.e., the first semester of 2023). The RF model exhibited the superior average AUC score. Therefore, the RF model that was built on the data from week 0 to week 4 in the source semester (i.e., Predictive Model 1 in Figure 1) was chosen for predicting at-risk students in week 5 in the target semester. The RF model that was built on the data from week 0 to week 7 in the source semester (i.e., Predictive Model 2 in Figure 1) was applied to predicting at-risk students in week 8 in the target semester.

3.2 At-Risk Student Identification

We applied RF models, developed during the training stage, to predict at-risk students at the start of weeks 5 and 8 in the target semester. Week 5 fell just before the midterm period (weeks 6–7) of the semester, which allowed the intervention to help prepare students for the upcoming first assessment. Coming after the midterm, week 8 was an ideal time for students to reflect on their performance based on the grades and feedback from the first assessment. Implementing an intervention in week 8 would support students in their reflective processes and aid in adjusting their learning approaches for the latter half of the semester.

In week 5, for example, we initially collected students' learning data in the past weeks, i.e., from week 0 to week 4, from the Ed and Moodle platforms. Subsequently, we developed features and inputted them in Predictive Model 1 to obtain the predicted performance for each student. Those students whose predicted performance indicated a "fail" were at-risk students. The same process was repeated to identify at-risk students in week 8.

3.3 Feedback Intervention

In both week 5 and week 8, we identified students at risk of failing. Immediately following each identification, we sent feedback emails (i.e., this was our intervention) to these at-risk students. An example of the intervention feedback emails is provided in the following.

Dear ⟨student name⟩, **[Sentence 1]**

Table 1. Features used in the predictive models. Each student interaction with the online learning platform at a specific time point was recorded as an event. A session is defined by the 85th percentile of the time gap between two consecutive events throughout the dataset. An active day refers to any day with at least one event related to learning actions (e.g., "forum contribution" or "course main page"). An active week is a week with a number of active days equal to or greater than the average number of active days per week observed among all students in the course. "LEARNING ACTION" refers to any of the six defined learning actions, namely, "forum contribution," "forum consumption," "grade report," "course main page," "learning materials," and "assignment."

Category	Feature	Description
General indicators	n_sessions_norm	Normalized session counts
	total_session_length	Total length of all student sessions
	total_session_length_norm	Normalized total length of all student sessions
	median_session_length_norm	Normalized median session length (in seconds)
	median_session_event_count_norm	Normalized median number of learning actions per session
	active_days_prop	Proportion of active days
	mdn_adays_prop	Proportion of median time distance between two consecutive active days
	active_weeks_prop	Proportion of active weeks
	avg_active_day_per_week	Average number of active days per week
Action-specific indicators	active_days_prop (<LEARNING ACTION>)	Proportion of active days for each learning action, such as "forum contribution"
	active_weeks_prop (<LEARNING ACTION>)	Proportion of active weeks for each learning action, such as "forum contribution"
Regularity of general indicators	session_length_entropy	Entropy of session length (in seconds), as a measure of (ir)regularity
	session_event_counts_entropy	Entropy of learning action counts per session, as a measure of (ir)regularity
	first_wday_prop	Proportion of weeks when the student was active on the first day of the week
	avg_1st_wday_ev_prop	Average proportion of learning actions on the first day of the week
Regularity of action-specific indicators	entropy_daily_event_count (<LEARNING ACTION>)	Entropy of daily action counts for each learning action, such as "forum contribution"
	entropy_weekly_event_count (<LEARNING ACTION>)	Entropy of weekly event counts for each learning action, such as "forum contribution"
Demographics	age	Student age
	gender	Male or female
	combined course	Degree type, single degree or double degree
	citizenship	Domestic, international, or other
	count_<GRADE>	Number of each final grade (i.e., HD, D, C, P, and N) the student obtained in previously completed courses
	commencement_experience	Number of years from the year the student started degree to the year enrolled in the course

We hope you have enjoyed learning this unit so far. **[Sentence 2]**
To succeed in this unit, students are expected to dedicate 12 hours a week in both online and offline learning activities. **[Sentence 3]**
You spent 0.43 hours on average per week learning this unit online so far. **[Sentence 4]**
Remember, there are several key learning activities in Ed Lesson. **[Sentence 5]**

In the past weeks, we expect you to complete 19 activities. **[Sentence 6]**
You have visited 0 of them. **[Sentence 7]**
Make sure you take time to complete all the activities. **[Sentence 8]**
Why not also take some time to think about the key takeaways and any questions that you may have and share with your peers in Ed Forum? **[Sentence 9]**
Here is a kind reminder to complete the remaining activities in Ed Lesson: **[Sentence 10]**
0. Week 01 - Introduction to Python, Variables, Statements and Expressions - W1 Pre-Workshop
1. Week 02 - Conditionals and Iteration (while loops) - W2 Applied
...
Don't forget to complete any remaining activities that you may still have in the previous weeks. **[Sentence 11]**
Please note that the observation of your online engagement is based on trace data between ⟨start time⟩ and ⟨end time⟩ only, which does not capture all the learning activities that you are involved in online and offline. **[Sentence 12]**
And remember, the teaching team is here to help you. **[Sentence 13]**
If you have any questions about this unit or face any learning challenges, you can reach out to the teaching team via Ed Forum and during our weekly teaching activities. **[Sentence 14]**
Do not forget that the next assignment will be due on ⟨date and time⟩. You can find the details about the assignment via ⟨link⟩. **[Sentence 15]**
If you do not wish to receive this feedback again, please let the teaching team know directly. **[Sentence 16]**

The feedback intervention was designed based on the framework of relational feedback (Dai et al., 2025). This framework was proposed to enhance feedback effectiveness by strengthening students' relationships with the learning community, such as instructors and peers. A detailed rationale for adopting the relational feedback framework as the foundation of our feedback intervention was elaborated on in Section 2.3. As suggested by Dai and colleagues (2025), feedback should (i) clarify students' performance (**Principle 1**), (ii) contain actionable information for students to improve their learning (**Principle 2**), (iii) encourage students to communicate with the learning community (**Principle 3**), and (iv) evoke students' positive emotions (**Principle 4**).

Guided by the relational feedback framework's four principles, four feedback researchers collaborated on the design of the feedback message. Before the message was drafted, a meeting was arranged with all the feedback researchers involved to decide on what specific information should be included in the feedback for each principle. Following the preliminary draft by one of the researchers, the feedback message underwent rigorous revisions by all members of our research and teaching team involved in the study to ensure that the tone was positive and supportive, aiming to motivate and uplift students. In the following, we map each sentence in the feedback to its corresponding relational principle and explain how it operationalizes that principle.

Principle 1: Clarifying student performance.

> **[Sentence 3]:** *To succeed in this unit, students are expected to dedicate 12 hours a week in both online and offline learning activities.*
>
> **[Sentence 4]:** *You spent 0.43 hours on average per week learning this unit online so far.*
>
> **[Sentence 6]:** *In the past weeks, we expect you to complete 19 activities.*
>
> **[Sentence 7]:** *You have visited 0 of them.*

We conducted a thorough analysis of the important features derived from predictive analysis. We identified a feature named "total_session_length" as a consistently top-ranking feature across all the predictive models built in the training stage. This feature was the total duration of student engagement with the online learning platform. The consistent importance made it a reliable metric to help students understand that putting effort into online learning would contribute to academic success. However, informing students of their total time spent could be overwhelming, and could demotivate them if they perceived themselves as lagging behind their goals or expectations. Therefore, the feedback was structured to report average weekly time spent calculated by dividing "total_session_length" by the number of weeks leading up to the intervention. Additionally, we established the weekly goal for time spent on both online and offline activities to enhance students' understanding of their current progress. Along with the time spent, the feedback also informed students of their progress in completing online learning activities to help them better identify areas requiring further attention.

Principle 2: Encouraging students' actions for improving learning engagement.

[Sentence 10]: *Here is a kind reminder to complete the remaining activities in Ed Lesson.*

[Sentence 11]: *Don't forget to complete any remaining activities that you may still have in the previous weeks.*

[Sentence 15]: *Do not forget that the next assignment will be due on ⟨date and time⟩. You can find the details about the assignment via ⟨link⟩.*

Table 2. Demographic information of students enrolled in the source and target semesters.

Demographic Categories	Source semester ($n = 1,003$)	Target semester ($n = 1,599$)
Gender		
Female	255 (25.42%)	464 (29.02%)
Male	746 (74.38%)	1,134 (70.92%)
Unknown	2 (0.20%)	1 (0.06%)
Age		
<18	0 (0.00%)	72 (4.50%)
18–24	972 (96.91%)	1,497 (93.62%)
25–34	28 (2.79%)	25 (1.56%)
35–44	3 (0.30%)	3 (0.19%)
45+	0 (0.00%)	2 (0.13%)
Years since degree commencement		
0	671 (66.90%)	1,139 (71.23%)
1	184 (18.34%)	275 (17.20%)
2	84 (8.37%)	105 (6.57%)
3	33 (3.29%)	54 (3.38%)
4+	31 (3.09%)	26 (1.63%)
Degree type		
Single	544 (54.24%)	883 (55.22%)
Double	438 (43.67%)	710 (44.40%)
Unknown	21 (2.09%)	6 (0.38%)
Citizenship		
Domestic	722 (71.98%)	1,193 (74.61%)
International	271 (27.02%)	406 (25.39%)
Others	10 (1.00%)	0 (0.00%)
High Distinction (HD)		
0	996 (99.30%)[1]	1,438 (89.93%)
1+	7 (0.70%)[2]	161 (10.07%)
Distinction (D)		
0	988 (98.50%)	1,338 (83.68%)
1+	15 (1.50%)	261 (16.32%)
Credit (C)		
0	989 (98.60%)	1,364 (85.30%)
1+	14 (1.40%)	235 (14.70%)
Pass (P)		
0	976 (97.31%)	1,404 (87.80%)
1+	27 (2.69%)	195 (12.20%)
Fail (N)		
0	999 (99.60%)	1,407 (87.99%)
1+	4 (0.40%)	192 (12.01%)

[1] 99.30% of students in the source semester had never achieved a High Distinction (HD) grade in any previously completed courses.

[2] 0.70% of students in the source semester achieved a High Distinction (HD) grade in at least one previously completed course.

The feedback message highlighted key learning activities on the Ed platform that had not been visited. Information about student visits to learning activities was retrieved from the Ed platform. This approach aimed to help students focus on the areas requiring more attention, thereby promoting more efficient learning by prioritizing essential tasks that deepen course comprehension. Knowing what has been completed and what remains could motivate students to take actions to stay on track with their coursework. Additionally, the feedback intervention reminded students of the deadline for the upcoming assignment and provided direct access to the assignment details, thereby motivating them to take timely action and engage with the task.

Principle 3: Inviting further communication.

> **[Sentence 9]:** *Why not also take some time to think about the key takeaways and any questions that you may have and share with your peers in Ed Forum?*
>
> **[Sentence 13]:** *And remember, the teaching team is here to help you.*
>
> **[Sentence 14]:** *If you have any questions about this unit or face any learning challenges, you can reach out to the teaching team via Ed Forum and during our weekly teaching activities.*

To invite students' further communication with peers and the teaching team, the feedback aimed to encourage students to share their learning outcomes and questions with peers and the teaching team.

Principle 4: Evoking students' positive emotions.

> **[Sentence 1]:** *Dear ⟨student name⟩*
>
> **[Sentence 2]:** *We hope you have enjoyed learning this unit so far.*
>
> **[Sentence 10]:** *Here is a kind reminder to complete the remaining activities in Ed Lesson.*
>
> **[Sentence 13]:** *And remember, the teaching team is here to help you.*
>
> **[Sentence 14]:** *If you have any questions about this unit or face any learning challenges, you can reach out to the teaching team via Ed Forum and during our weekly teaching activities.*

These sentences evoke positive emotions by creating a sense of care, support, and encouragement. Personalizing the message with the student's name and the salutation "Dear" establishes a friendly and respectful tone. Expressing hope that the student has enjoyed the unit shows empathy and promotes a positive reflection on their learning experience. Phrasing the reminder as "kind" conveys supportiveness rather than making it feel like a demand. Finally, reassuring the student that the teaching team is there to help reinforces a sense of belonging and reduces anxiety.

3.4 RQ1: Evaluation of the Generalizability of the Predictive Model

Although our predictive models (i.e., Predictive Models 1 and 2, developed based on RF) were applied to identify at-risk students in a new semester of the same course and the models were initially trained on the data from the offering of the course in the previous semester, variations in the cohorts between two semesters (as illustrated in Table 2) could potentially impact model performance. To assess the generalizability of the predictive models used in our study, we conducted a post-evaluation after the conclusion of the target semester. We evaluated the prediction performance of the RF model (details on the selection process of the model can be seen in Section 3.1.2) across three different sets of training and test datasets created based on the data from week 0 to week 4, referred to as "Model 1 (0–4)," "Model 2 (0–4)" (i.e., Predictive Model 1 in Figure 1), and "Model 3 (0–4)," as shown in Table 3. Model 1 (0–4) was trained and tested on the data from week 0 to week 4 in the source semester. This model acted as a baseline to establish the performance of the predictive model that was trained and tested on the source data. Model 2 (0–4) was trained on the data from week 0 to week 4 in the source semester and tested on the data from week 0 to week 4 in the target semester. This model corresponds to Predictive Model 1 in Figure 1, which we used to identify at-risk students in week 5 in the target semester. Model 3 (0–4) was trained and tested on the data from week 0 to week 4 in the target semester. This model was used to assess how well the RF model could perform when deployed directly in the new context (i.e., both trained and tested in the target semester) instead of trained in the source semester and tested in the target semester. The RF model was also evaluated across three different sets of training and test datasets using the data from week 0 to week 7, which also leads to three models, i.e., "Model 1 (0–7)," "Model 2 (0–7)," and "Model 3 (0–7)," as shown in Table 4. Specifically, Model 2 (0–7) corresponds to Predictive Model 2 in Figure 1, which we applied in week 8 in the target semester.

The prediction performance was evaluated using accuracy, macro F1 score, and AUC score. In addition to these widely used metrics, we included two additional measurements: negative predictive value (NPV) and true negative rate (TNR). NPV is the proportion of students who actually failed the course among those identified as at risk. TNR is the proportion of students

Table 3. The statistical difference in predictive performance between Model 1 (0–4), Model 2 (0–4), and Model 3 (0–4), measured by ANOVA tests.

	Week 0–Week 4					
	Model 1 (0–4)	Model 2 (0–4)	Model 3 (0–4)			
Train	Source	Source	Target	F-statistic	p-value	Eta squared
Test	Source	Target	Target			
Accuracy	0.93	0.82	0.91	54.75	9.29e-07	0.90
Macro F1 score	0.93	0.63	0.91	312.45	4.47e-11	0.98
AUC score	0.98	0.72	0.96	826.63	1.40e-13	0.99
NPV	0.94	0.42	0.92	560.11	1.42e-12	0.99
TNR	0.92	0.35	0.88	400.57	1.03e-11	0.99

Table 4. The statistical difference in predictive performance between Model 1 (0–7), Model 2 (0–7), and Model 3 (0–7), measured by ANOVA tests.

	Week 0–Week 7					
	Model 1 (0–7)	Model 2 (0–7)	Model 3 (0–7)			
Train	Source	Source	Target	f-value	p-value	Eta squared
Test	Source	Target	Target			
Accuracy	0.92	0.84	0.91	46.28	2.28e-06	0.89
Macro F1 score	0.92	0.68	0.91	338.20	2.81e-11	0.98
AUC score	0.98	0.78	0.96	820.53	1.46e-13	0.99
NPV	0.93	0.49	0.92	307.35	4.83e-11	0.98
TNR	0.92	0.45	0.90	228.03	2.84e-10	0.97

who were correctly identified as at risk among those who failed the course. High NPV and TNR indicate that the model could accurately detect students who were at risk (i.e., true negative instances), which was the goal of our study. NPV also offers critical insights into the proportion of students who although not at risk, may be bothered by interventions. To ensure the reliability of our comparison, we employed five-fold cross-validation for each model. We then conducted ANOVA tests to determine if there were significant differences in performance across the models, with subsequent t-tests as post hoc tests to pinpoint specific pairs of models between which the differences were statistically significant. To control for the increased risk of type I errors due to multiple comparisons, we applied Bonferroni correction to the p-values of these t-tests.

To uncover key factors influencing student success across different semesters, and to inform future interventions, further analysis was carried out on the important features contributing to the predictive power of Model 1 (0–4) and Model 3 (0–4), as well as between Model 1 (0–7) and Model 3 (0–7). Specifically, we quantified the contribution of each feature to the predictive models using SHAP (SHapley Additive exPlanations) values (Lundberg & Lee, 2017) and extracted the top 10 important features for comparison.

3.5 RQ2: Evaluation of Student Engagement Change after Feedback Intervention

In the feedback intervention given to at-risk students, we listed the learning activities that had not been visited by the students on the Ed platform before the week of intervention. To evaluate the effectiveness of this intervention, for each at-risk student, we examined whether the student engaged with the unvisited learning activities listed in the provided feedback message. To understand the engagement change for individual students, we calculated descriptive statistics for the number of unique activities visited by at-risk students before the intervention (#Before), one week post (#One Week), and two weeks post intervention (#Two Weeks). Statistics include the mean (Mean), standard deviation (Std), minimum (Min), 25th percentile (25%), median

(Median), 75th percentile (75%), and maximum (Max). Paired t-tests were conducted to investigate the significance of the differences among three groups, i.e., #Before vs. #One Week, #Before vs. #Two Weeks, and #One Week vs. #Two Weeks. Bonferroni correction was applied to the p-values of these t-tests. We also calculated the following four metrics to compare overall students' engagement with learning activities between the time before the intervention and the time one week and two weeks after the intervention: (i) the number of at-risk students who visited at least one originally unvisited learning activity and their proportion relative to all at-risk students, (ii) the average number of learning activities visited by at-risk students, (iii) the average amount of time in hours that at-risk students spent on all the required learning activities, and (iv) the average amount of time in hours that at-risk students spent on originally unvisited learning activities.

3.6 RQ3: Evaluation of Student Perceptions of Feedback Intervention

To further understand the impact of our feedback intervention, we designed a survey and distributed it to all the at-risk students identified in either week 5 or week 8 of the target semester (the survey is available at this link). The survey was delivered via email and was structured into two sections. The first part contained six closed questions to collect students' background information, such as their gender, year of study, and first language background. The second part consisted of 10 five-point Likert scale questions to ask students' perceptions on the feedback intervention and one multiple-choice question to investigate students' preferences on the frequency of receiving such intervention. The 10 five-point Likert scale questions probed aspects such as the overall satisfaction, whether the feedback is understandable and timely, how students responded to the feedback emotionally, how the feedback helped students improve their learning, and which specific aspects of feedback motivated students to act.

After gathering responses from students, we computed the descriptive statistics for the scores assigned to each of the 10 five-point Likert scale questions. These statistics include the mean (Mean), median (Median), and standard deviation (Std). Regarding students' preferences for the frequency of feedback intervention, we quantified the number of students across the different options provided in the multiple-choice question.

4. Results

4.1 Results on RQ1

For the models trained and tested on the data from week 0 to week 4 shown in Table 3, Model 1 (0–4) showed the highest accuracy (0.93), indicating excellent performance when both trained and tested within the same semester. Model 3 (0–4), which was trained and tested within the target semester, also demonstrated strong accuracy (0.91). In contrast, Model 2 (0–4), which was trained on the source semester and tested on the target semester, showed considerably lower accuracy (0.82). Similar trends were observed with the macro F1 score, AUC score, NPV, and TNR. Statistical tests confirmed significant differences in model performance across different training and testing conditions. Post-hoc t-tests pinpointed these differences specifically between Models 1 (0–4) and 2 (0–4), and Models 2 (0–4) and 3 (0–4), for all metrics (see Table 5 for details). For the models trained and tested on the data from week 0 to week 7 (Table 4), consistent with the findings in Tables 3 and 5, Model 1 (0–7) and Model 3 (0–7) were superior to Model 2 (0–7) on the five performance measurements. Post-hoc t-tests again confirmed these differences between Models 1 (0–7) and 2 (0–7), and between Models 2 (0–7) and 3 (0–7), for all metrics (see Table 6).

The results in Tables 3 and 4 suggest that models trained and tested within the same semester (Model 1 and Model 3) significantly outperform models trained on past semester data and tested on current semester data (Model 2). However, the inferior performance of Model 2 does not imply that the model derived from the original dataset is incapable of making predictions on a new dataset. On the contrary, Model 2, which was constructed using a dataset from either weeks 0 to 4 or weeks 0 to 7, demonstrated quite satisfactory performance, achieving AUC scores above 0.7. This suggests that the model built in the source semester exhibited strong generalizability when applied to a new semester.

To further understand the differences between the predictive models developed for the source semester and the target semester, we analyzed feature importance in both models. Figures 2 and 3 illustrate the top 10 important features impacting the predictions of Model 1 and Model 3 in Tables 3 and 4, respectively. We found that all the top 10 features are related to student engagement rather than student demographics, indicating that levels of student engagement are more crucial in determining academic success. In Figure 2, Model 1 (0–4) and Model 3 (0–4) have eight features in common (marked by *) among the top 10 important features. Six out of eight common features are general indicators of student engagement with the online learning platforms. Similar results can be seen from the comparison between the top 10 important features in Model 1 (0–7) and Model 3 (0–7), as shown in Figure 3. Seven out of 10 features contributed to the performance of both Model 1 (0–7) and Model 3 (0–7), and the majority are general engagement predictors. The substantial number of important features shared in both models provides strong support for the generalizability of the model that we built in the source semester to predict at-risk students in the target semester.

Although models built in various semesters share many important features, there is a notable difference in their rankings among the top 10 critical features. For example, in Figure 2, features associated with session length, specifically

Table 5. The statistical differences in predictive performance between model pairs, assessed using post hoc t-tests. Models were constructed using data from week 0 to week 4. Significant adjusted p-values are denoted by a *.

Week 0–Week 4			
Groups	**t-score (df=8)**	**adjusted p-value[1]**	**Cohen's d**
Accuracy			
Model 1 (0–4) vs. Model 2 (0–4)	9.22	4.63e-05*	5.84
Model 1 (0–4) vs. Model 3 (0–4)	1.55	0.48	0.98
Model 2 (0–4) vs. Model 3 (0–4)	-11.31	1.01e-05*	-7.16
Macro F1 score			
Model 1 (0–4) vs. Model 2 (0–4)	21.15	7.86e-08*	13.38
Model 1 (0–4) vs. Model 3 (0–4)	1.54	0.48	0.98
Model 2 (0–4) vs. Model 3 (0–4)	-29.49	5.68e-09*	-18.65
AUC score			
Model 1 (0–4) vs. Model 2 (0–4)	32.04	2.94e-09*	20.27
Model 1 (0–4) vs. Model 3 (0–4)	2.05	0.22	1.30
Model 2 (0–4) vs. Model 3 (0–4)	-37.33	8.74e-10*	-23.61
NPV			
Model 1 (0–4) vs. Model 2 (0–4)	26.55	1.31e-08*	16.79
Model 1 (0–4) vs. Model 3 (0–4)	0.88	1.22	0.56
Model 2 (0–4) vs. Model 3 (0–4)	-28.51	7.42e-09*	-18.03
TNR			
Model 1 (0–4) vs. Model 2 (0–4)	23.25	3.73e-08*	14.71
Model 1 (0–4) vs. Model 3 (0–4)	1.53	0.49	0.97
Model 2 (0–4) vs. Model 3 (0–4)	-28.53	7.38e-09*	-18.05

[1] p-value was adjusted using Bonferroni correction.

"total_session_length" and "total_session_length_norm," were ranked second and third in Model 1 (0–4), respectively. Conversely, in Model 3 (0–4), these features descended to the fifth and sixth positions. Instead, features related to the proportion of active days, namely "active_days_prop" and "active_days_prop (learning materials)" in Model 1 (0–4), ascended to the second and third ranks in Model 3 (0–4). This change may indicate distinct behavioural patterns of students between the source and target semesters, which could explain the decrease of the prediction accuracy of Model 2 (0–4).

4.2 Results on RQ2

In week 5 of the target semester, 239 students were identified as at risk. As shown in Table 7, prior to the feedback intervention in week 5, these students had visited fewer than three (2.69 on average) out of 19 required activities. One week following the intervention, this average increased to 3.66, and further rose to 3.92 two weeks after the intervention. The second feedback intervention also resulted in an increase in the number of activities visited by students. Specifically, before the intervention, at-risk students visited approximately five activities on average. Two weeks following the intervention, this number increased to more than six. The increases in the number of unique activities visited post intervention were statistically significant, as demonstrated by the t-test results presented in Table 8.

As illustrated in Table 9, one week after the first feedback intervention, 58 at-risk students engaged with at least one previously unvisited learning activity, representing 24.27% of all 239 at-risk students identified in week 5. This number increased to 76, accounting for 31.80% of all at-risk students two weeks post intervention. Concurrently, the engagement with

Table 6. The statistical differences in predictive performance between model pairs, assessed using post hoc t-tests. Models were constructed using data from week 0 to week 7. Significant adjusted p-values are denoted by a *.

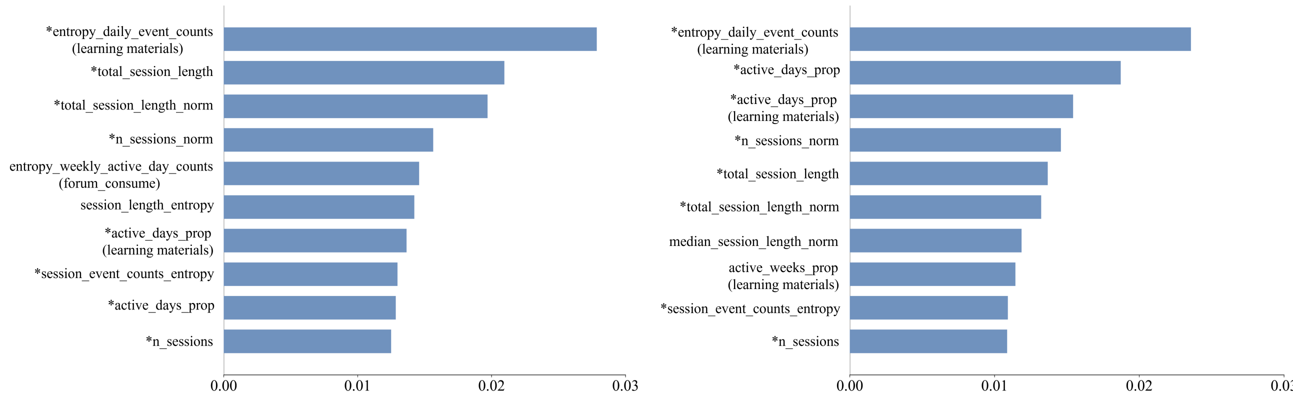
	Week 0–Week 7		
Groups	**t-score**	**adjusted p-value[1]**	**Cohen's d**
Accuracy			
Model 1 (0–7) vs. Model 2 (0–7)	8.17	1.13e-04*	5.16
Model 1 (0–7) vs. Model 3 (0–7)	1.11	0.90	0.70
Model 2 (0–7) vs. Model 3 (0–7)	-9.00	5.56e-03*	-5.69
Macro F1 score			
Model 1 (0–7) vs. Model 2 (0–7)	22.23	5.32e-08*	14.06
Model 1 (0–7) vs. Model 3 (0–7)	1.11	0.90	0.70
Model 2 (0–7) vs. Model 3 (0–7)	-30.48	4.37e-09*	-19.28
AUC score			
Model 1 (0–7) vs. Model 2 (0–7)	46.11	1.62e-10*	29.16
Model 1 (0–7) vs. Model 3 (0–7)	2.53	0.11	1.60
Model 2 (0–7) vs. Model 3 (0–7)	-32.58	2.58e-09*	-20.60
NPV			
Model 1 (0–7) vs. Model 2 (0–7)	18.18	2.59e-07*	11.50
Model 1 (0–7) vs. Model 3 (0–7)	1.12	0.88	0.71
Model 2 (0–7) vs. Model 3 (0–7)	-19.85	1.30e-07*	-12.55
TNR			
Model 1 (0–7) vs. Model 2 (0–7)	16.64	5.16e-07*	10.52
Model 1 (0–7) vs. Model 3 (0–7)	0.72	1.48	0.45
Model 2 (0–7) vs. Model 3 (0–7)	-17.21	3.96e-07*	-10.89

[1] p-value was adjusted using Bonferroni correction.

new learning activities contributed to an increase in the average number of learning activities visited by at-risk students, rising from 2.69 prior to the intervention to 3.66 one week after the intervention and 3.92 two weeks following the intervention. The increased engagement with learning activities was also evident from the average time that at-risk students spent on all required learning activities, which was 0.95 hours before, 1.21 hours one week after, and 1.25 hours two weeks after the intervention. Of the 0.26 additional hours observed one week post intervention (calculated by subtracting 0.95 from 1.21), 0.10 hours were devoted to activities that had not been visited previously. Two weeks after the intervention, on average, at-risk students spent 1.25 hours on originally unvisited learning activities.

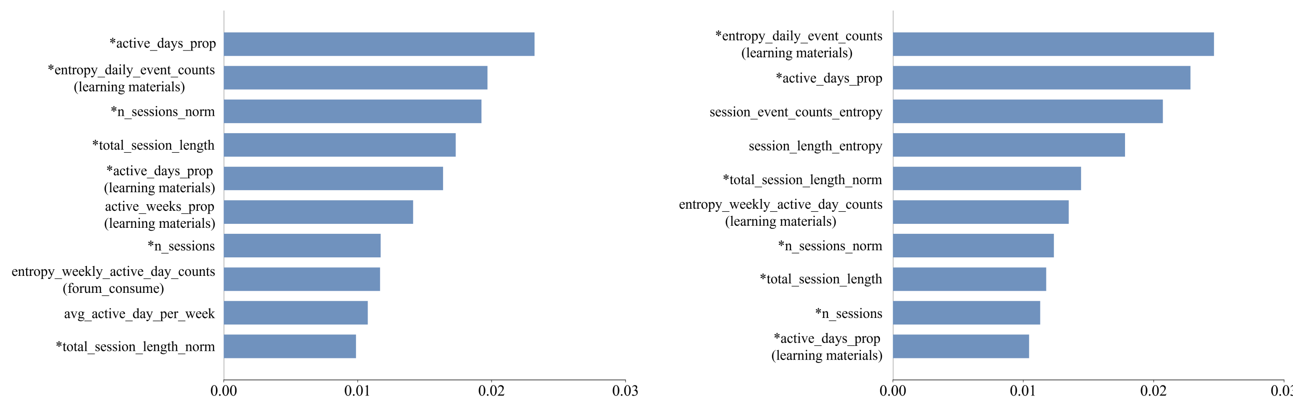
The changes in engagement with learning activities for 252 at-risk students identified in week 8 are detailed in Table 10. One week following the second feedback intervention, 65 of these students visited learning activities that they had not engaged with previously. This number rose to 96 two weeks post intervention, accounting for 38.10% of all at-risk students. The average number of learning activities visited by at-risk students increased from 5.04 before the intervention to 5.45 one week after and to 6.38 two weeks after the intervention. In terms of time spent, at-risk students dedicated an average of 1.62 hours to all assigned learning activities before the intervention, which increased to 1.70 hours one week after and to 1.75 hours two weeks after the intervention. Specifically, on average, these students spent 0.05 hours on originally unvisited learning activities one week after the intervention, which rose to 0.10 hours two weeks post intervention.

Figure 2. Top 10 most important features of Model 1 and Model 3 in Table 3 measured by SHAP values. Common features are marked by *.



(a) Model 1 in Table 3 **(b)** Model 3 in Table 3

Figure 3. Top 10 most important features of Model 1 and Model 3 in Table 4 measured by SHAP values. Common features are marked by *.



(a) Model 1 in Table 4 **(b)** Model 3 in Table 4

Table 7. Descriptive statistics on the number of unique activities visited by at-risk students before (#Before), one week after (#One Week), and two weeks after (#Two Weeks) the feedback intervention. #Assigned denotes the number of unique activities assigned to students before intervention.

	First feedback intervention				Second feedback intervention			
	#Assigned	**#Before**	**#One Week**	**#Two Weeks**	**#Assigned**	**#Before**	**#One Week**	**#Two Weeks**
Mean	19	2.69	3.66	3.92	33	5.04	5.45	6.38
Std	0	4.31	5.10	5.18	0	7.24	7.48	8.09
Min	19	0	0	0	33	0	0	0
25%	19	0	0	0	33	0	0	0
Median	19	1	1	2	33	2	2	3
75%	19	3	5	6	33	7	8	9
Max	19	19	19	19	33	33	33	33

4.3 Results on RQ3

Out of 356 students who received the survey (i.e., the students who received our email feedback), we obtained 33 responses, which is comparable to the commonly observed response rate to surveys in learning analytics (Whitelock-Wainwright et al.,

Table 8. The statistical difference between the number of unique activities at-risk students visited before, one week after, and two weeks after the feedback intervention, assessed using t-tests. Significant values are denoted by a *.

Groups	t-score	df	adjusted p-value[1]	Cohen's d
First feedback intervention				
#Before vs. #One Week	-6.43	238	2.09e-09*	-0.20
#Before vs. #Two Weeks	-7.51	238	3.46e-12*	-0.26
#One Week vs. #Two Weeks	-4.34	238	6.41e-05*	-0.05
Second feedback intervention				
#Before vs. #One Week	-8.01	251	1.31e-13*	-0.06
#Before vs. #Two Weeks	-7.07	251	4.51e-11*	-0.17
#One Week vs. #Two Weeks	-5.23	251	1.06e-06*	-0.12

[1]p-value was adjusted using Bonferroni correction.

Table 9. The change in engagement with learning activities among 239 at-risk students, identified in week 5, following the first feedback intervention.

First feedback intervention			
	Before intervention	**One week post intervention**	**Two weeks post intervention**
Number of at-risk students who visited at least one originally unvisited learning activity (and their proportion relative to all detected at-risk students)	N/A	58 (24.27%)	76 (31.80%)
Average number of learning activities visited by at-risk students	2.69	3.66	3.92
Average amount of time in hours that at-risk students spent on all the required learning activities	0.95	1.21	1.25
Average amount of time in hours that at-risk students spent on originally unvisited learning activities	N/A	0.10	0.13

2020). To assess the internal consistency of the survey, we calculated Cronbach's α, which yielded a value of $\alpha = 0.985$. In this study, internal consistency refers to the degree to which the survey items consistently measure students' perceptions of the feedback intervention, specifically its clarity, usefulness, emotional impact, and motivational effect. This high alpha value suggests that the items are highly interrelated and likely reflect a single underlying construct, i.e., students' overall response to the feedback intervention. The mean and median of students' responses to 10 five-point Likert scale questions are presented in Table 11. The average responses to all survey questions exceeded 3 points (indicating "Neutral"), which suggests that students generally held positive attitudes toward the feedback email. Analysis of the median values reveals that most at-risk students agreed that the provided feedback intervention was timely (Question 6), was easily understandable (Question 2), made them feel cared for (Question 4), and helped them develop and adjust their learning strategies (Question 8). Regarding the specific aspects of the feedback that motivated students to take action (Question 10), responses highlighted the clarity of feedback, timeliness, personalization, information about their current progress, and details about unvisited activities as particularly effective, each earning a median score of 4 (indicating "Agree").

Figure 4 illustrates student preferences regarding the frequency of the feedback intervention. It is evident that the majority of students prefer receiving feedback more frequently than the twice-a-semester frequency implemented in the target semester.

Table 10. The change in engagement with learning activities among 252 at-risk students, identified in week 8, following the second feedback intervention.

Second feedback intervention			
	Before intervention	**One week post intervention**	**Two weeks post intervention**
Number of at-risk students who visited at least one originally unvisited learning activity (and their proportion relative to all detected at-risk students)	N/A	65 (25.79%)	96 (38.10%)
Average number of learning activities visited by at-risk students	5.04	5.45	6.38
Average amount of time in hours that at-risk students spent on all the required learning activities	1.62	1.70	1.75
Average amount of time in hours that at-risk students spent on originally unvisited learning activities	N/A	0.05	0.10

5. Discussion

We conducted a study that evaluated the effectiveness of a learning analytics–based approach for automated detection and support of at-risk students at the early stage of a course. The approach could be relatively easily scaled across multiple courses as the features used for training predictive models are course agnostic, making it a relatively cost effective method for large-scale educational settings. Such an automatic approach could alleviate instructors' workload in monitoring student learning progress and performing timely and relational interventions. Students who know that the institution cares about students' learning progress would be motivated to engage more deeply with course materials and participate more actively in learning activities.

5.1 Interpretation of the Results

5.1.1 RQ1

The predictive models deployed in the target semester to identify at-risk students early were trained using data from a previous semester. We observed a significant decrease in prediction accuracy when these models were applied to a new dataset, which is consistent with findings reported in López-Zambrano and colleagues (2020). However, López-Zambrano and colleagues (2020) reported that the performance of predictive models applied to new student cohorts was approximately equivalent to random chance, as indicated by AUC scores. In contrast, our predictive models, when applied to a different student cohort, achieved AUC scores above 0.7. A possible explanation for this discrepancy could be that the dataset used to train and test models by López-Zambrano and colleagues (2020) originated from different courses, whereas in our study, both the source dataset for training and the target dataset for testing were derived from different offerings of the same course, under very similar instructional conditions. As highlighted by Gašević and colleagues (2016) and López-Zambrano and colleagues (2020), the similarities in the course instructional conditions—such as the teaching team, the number and design of assessments, and the use of learning platforms—likely contribute to the generalizability of the predictive model. Furthermore, the similarity in learning behaviours between different cohorts of students might also explain the high accuracy of our models. According to our analysis, predictive models built in the source and target semesters shared a substantial number of important features related to students' interactions with online learning platforms.

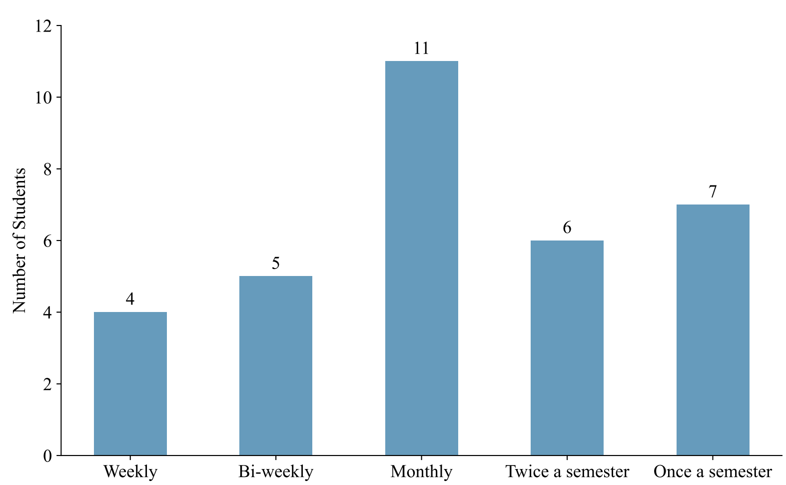
5.1.2 RQ2

The literature (R. Baker et al., 2016; Borrella et al., 2019, 2022) indicates that email-based interventions have not shown a significantly positive impact on overall student retention. In contrast, by examining their effects on individual students, our study found that at-risk students significantly increased their engagement with learning activities following the receipt of feedback interventions. Despite not employing randomized experiments, our study shows that the observed increase in engagement after intervention indicates a potentially positive influence of our feedback intervention on student learning behaviours, such as engaging with previously unvisited learning activities. The change in students' learning behaviours may contribute to an enhancement in their overall learning performance, thereby reducing the possibility of dropping out of their study (R. Baker

Table 11. The mean (Mean), median (Median), and standard deviation (Std) of students' responses to the 10 five-point Likert scale survey questions.

No.	Question focus	Mean	Median	Std
1	Satisfaction with feedback	3.30	3	1.16
2	Feedback understandability	3.82	4	1.24
3	Helped reflect on engagement	3.18	3	1.18
4	Felt cared for by teaching team	3.52	4	1.37
5	Feeling about learning progress	3.27	3	1.10
6	Timeliness of feedback	3.85	4	1.03
7	Motivated time allocation for learning	3.27	3	1.18
8	Helped develop learning strategies	3.24	4	1.25
9	Helped improve time management	3.18	3	1.21
10	Rate how much each feedback aspect motivated you to act:			
	(1) Clarity of the feedback	3.48	4	1.18
	(2) Relevance of the feedback to the learning goals	3.30	3	1.24
	(3) Timeliness of the feedback	3.55	4	1.09
	(4) Personalization of the feedback	3.52	4	1.15
	(5) Information about the current progress	3.64	4	1.29
	(6) Information about remaining activities	3.48	4	1.25

Figure 4. Students' preference about the frequency of the feedback intervention.



et al., 2016).

5.1.3 RQ3

Our investigation into RQ3 contributes novel insights by specifically examining the subjective perceptions of at-risk students on feedback intervention, an area that has received limited attention in prior research. The positive perceptions revealed through our qualitative study, complemented by the significant increase in engagement with learning activities observed in our quantitative experiments, underscore the usefulness of our feedback intervention in encouraging student participation in learning. Furthermore, by focusing on individual experiences and perceived impacts, rather than solely quantitative outcomes, we provide a deeper understanding of how feedback interventions can be tailored to effectively support at-risk students. According to our findings, specific aspects of our feedback, such as the timeliness of the feedback and the provision of information about previously unvisited activities, were particularly motivational for students, suggesting pathways for further refinement of the feedback intervention. For instance, students who responded to the survey expressed different preferences regarding the frequency of feedback intervention. Given that timeliness is closely associated with the impact of the feedback intervention (Poulos & Mahony, 2008), tailoring the delivery of intervention to better align with individual student expectations could enhance feedback effectiveness. To this end, allowing students to request feedback intervention on demand, rather than adhering to a uniform schedule, might be a more adaptive and effective approach.

5.2 Implications

Although our predictive models achieved a satisfactory performance in identifying at-risk students, the observed decrease in the prediction performance signals a need for improving the model using data from the new semester to maintain its accuracy across different student cohorts. For example, in the scenario in this study, i.e., early identification of at-risk students who might fail the course, where the final performance is not yet available at the early stage of the course, the model could be retrained using mid-semester test or assignment scores as proxies for final outcomes. This approach could enable the model to capture the learning behaviours of the new student cohort.

Combined with interventions, our predictive models could be incorporated in learning platforms for automatically identifying and supporting at-risk students. Specifically, learning platforms can monitor students' responses to interventions by evaluating and reporting changes in student engagement levels post intervention. Based on this information, educators can dynamically adjust to individual student needs, offering personalized guidance and support precisely when it is most needed. For instance, if a platform suggests that a student has not engaged with any learning activities one or two weeks after receiving email-based interventions, it could remind educators to offer more direct intervention strategies, such as scheduling face-to-face meetings, to ask students for their difficulties and encourage their participation in learning. This adaptive approach not only can aid in maintaining student engagement but also contributes to creating a relational learning environment where students feel that their learning progress is valued by the teaching staff and the institution. Such positive emotions can strengthen students' relationships with the learning community, ultimately leading to improved learning outcomes (Middleton et al., 2023).

5.3 Limitations and Future Work

Our predictive models have demonstrated potential in the early identification of at-risk students but could benefit from the incorporation of more granular features about student interactions with online learning platforms. Specifically, a detailed tracking of time spent on each required learning task; the identification of periods of inactivity or peak activity times; and monitoring the frequency of student log-ins, both daily and weekly, can provide deeper insights into student engagement patterns. Additionally, while the features used in our predictive models were derived from the literature, future work should explore data-driven approaches. For example, applying principle component analysis (PCA) could help develop latent features that summarize complex behaviour patterns, and employing feature importance analysis could aid in filtering out less informative features. These methods could not only improve the accuracy of our models but also enable the application of prescriptive analytics to suggest specific actions aimed at increasing learning outcomes. Based on these suggestions, we can design more actionable feedback that encourages student learning engagement (Liang et al., 2024).

Our evaluation of the impact of the feedback intervention did not account for additional variables that could influence student behaviour, such as individual motivation; student levels of skills to effectively self-regulate learning; and students' engagement outside of the Ed platform, e.g., attending an applied class. For example, while learning activities on the Ed platform were expected to be completed within the week they were assigned, some students may habitually complete all activities intensively at a specific time beyond the required week. Therefore, our measurement of student engagement with learning activities one and two weeks after the feedback intervention may not accurately reflect the actual impact of the intervention, potentially leading to an underestimation of its effectiveness. A more accurate method to assess the feedback effectiveness is through a randomized control experiment (Borrella et al., 2022), which would involve randomly assigning students to equally sized control and treatment groups. Only students in the treatment group would receive the feedback intervention. This experimental design could potentially mitigate the influence of external factors on measuring the impact of feedback intervention on student learning engagement by accounting for natural engagement changes over time (e.g., those caused by exam periods or course pacing) and more clearly distinguishing changes that are likely attributable to the intervention.

6. Conclusion

To enhance students' academic success, we conducted a study to identify at-risk students early by using trace data on the online learning platforms and deliver timely feedback interventions designed by feedback experts according to the theory of relational feedback. According to the evaluation after the completion of the semester, the predictive models developed from data collected in the previous semester demonstrated an acceptable ability to identify at-risk students at the early stage of the semester, illustrating their generalizability in different educational contexts. At-risk students who received the feedback intervention engaged more with their learning activities, which indicates the effectiveness of the feedback intervention. Students' positive perceptions of the provided feedback highlighted its usefulness in influencing students' learning behaviours.

Declaration of Conflicting Interest

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The authors declared no financial support for the research, authorship, and/or publication of this article

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. https://doi.org/10.1109/ACCESS.2021.3049446

Akçapınar, G., Altun, A., & Aşkar, P. (2019). Using learning analytics to develop early-warning system for at-risk students. *International Journal of Educational Technology in Higher Education*, *16*(1), 1–20. https://doi.org/10.1186/s41239-019-0172-z

Arnold, K. E., & Pistilli, M. D. (2012). Course signals at Purdue: Using learning analytics to increase student success. In *Proceedings of the Second International Conference on Learning Analytics and Knowledge* (LAK 2012), 29 April–2 May 2012, Vancouver, British Columbia, Canada (pp. 267–270). ACM. https://doi.org/10.1145/2330601.2330666

Bainbridge, J., Melitski, J., Zahradnik, A., Lauría, E. J., 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. https://doi.org/10.1080/15236803.2015.12001831

Baker, R., Evans, B., & Dee, T. (2016). A randomized experiment testing the efficacy of a scheduling nudge in a Massive Open Online Course (MOOC). *AERA Open*, *2*(4), 2332858416674007. https://doi.org/10.1177/2332858416674007

Baker, R. S., Lindrum, D., Lindrum, M. J., & Perkowski, D. (2015). Analyzing early at-risk factors in higher education e-learning courses. In O. Santos, C. Romero, M. Pechenizkiy, A. Merceron, P. Mitros, J. Luna, C. Mihaescu, P. Moreno, A. Hershkovitz, S. Ventura, & M. Desmarais (Eds.), *Proceedings of the Eighth International Conference on Educational Data Mining* (EDM 2015), 26–29 June 2015, Madrid, Spain (pp. 150–155). International Educational Data Mining Society. https://files.eric.ed.gov/fulltext/ED560503.pdf

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. https://doi.org/10.3390/app10134427

Berka, P., & Marek, L. (2021). Bachelor's degree student dropouts: Who tend to stay and who tend to leave? *Studies in Educational Evaluation*, *70*, 100999. https://doi.org/10.1016/j.stueduc.2021.100999

Borrella, I., Caballero-Caballero, S., & Ponce-Cueto, E. (2019). Predict and intervene: Addressing the dropout problem in a MOOC-based program. In *Proceedings of the Sixth (2019) ACM Conference on Learning @ Scale* (L@S 2019), 24–25 June 2019, Chicago, Illinois, USA (pp. 1–9). ACM. https://doi.org/10.1145/3330430.3333634

Borrella, I., Caballero-Caballero, S., & Ponce-Cueto, E. (2022). Taking action to reduce dropout in MOOCs: Tested interventions. *Computers & Education*, *179*, 104412. https://doi.org/10.1016/j.compedu.2021.104412

Brooks, C., Thompson, C., & Teasley, S. (2015). A time series interaction analysis method for building predictive models of learners using log data. In *Proceedings of the Fifth International Conference on Learning Analytics and Knowledge* (LAK 2015), 16–20 April 2015, Poughkeepsie, New York, USA (pp. 126–135). ACM. https://doi.org/10.1145/2723576.2723581

Campbell, J. P., DeBlois, P. B., & Oblinger, D. G. (2007). Academic analytics: A new tool for a new era. *EDUCAUSE review*, *42*(4), 40. https://er.educause.edu/articles/2007/7/academic-analytics-a-new-tool-for-a-new-era

Chipchase, L., Davidson, M., Blackstock, F., Bye, R., Clothier, P., Klupp, N., Nickson, W., Turner, D., & Williams, M. (2017). Conceptualising and measuring student disengagement in higher education: A synthesis of the literature. *International Journal of Higher Education*, *6*(2), 31–42. https://doi.org/10.5430/ijhe.v6n2p31

Christou, V., Tsoulos, I., Loupas, V., Tzallas, A. T., Gogos, C., Karvelis, P. S., Antoniadis, N., Glavas, E., & Giannakeas, N. (2023). Performance and early drop prediction for higher education students using machine learning. *Expert Systems with Applications*, *225*, 120079. https://doi.org/10.1016/j.eswa.2023.120079

Clow, D. (2012). The learning analytics cycle: Closing the loop effectively. In *Proceedings of the Second International Conference on Learning Analytics and Knowledge* (LAK 2012), 29 April–2 May 2012, Vancouver, British Columbia, Canada (pp. 134–138). ACM. https://doi.org/10.1145/2330601.2330636

Dai, W., Tsai, Y.-S., Gašević, D., & Chen, G. (2025). Designing relational feedback: A rapid review and qualitative synthesis. *Assessment & Evaluation in Higher Education*, *50*(1), 16–30. https://doi.org/10.1080/02602938.2024.2361166

Dawson, S., Jovanovic, J., Gašević, D., & Pardo, A. (2017). From prediction to impact: Evaluation of a learning analytics retention program. In *Proceedings of the Seventh International Conference on Learning Analytics and Knowledge* (LAK 2017), 13–17 March 2017, Vancouver, British Columbia, Canada (pp. 474–478). ACM. https://doi.org/10.1145/3027385.3027405

Dix, N., Lail, A., Birnbaum, M., & Paris, J. (2020). Exploring the "at-risk" student label through the perspectives of higher education professionals. *The Qualitative Report*, *25*(11). https://doi.org/10.46743/2160-3715/2020.3371

Er, E. (2012). Identifying at-risk students using machine learning techniques: A case study with IS 100. *International Journal of Machine Learning and Computing*, *2*(4), 476. https://www.ijml.org/show-32-132-1.html

Faas, C., Benson, M. J., Kaestle, C. E., & Savla, J. (2018). Socioeconomic success and mental health profiles of young adults who drop out of college. *Journal of Youth Studies*, *21*(5), 669–686. https://doi.org/10.1080/13676261.2017.1406598

Falkner, N. J., & Falkner, K. E. (2012). A fast measure for identifying at-risk students in computer science. In *Proceedings of the Ninth Annual International Conference on International Computing Education Research* (ICER 2012), 9–11 September 2012, Auckland, New Zealand (pp. 55–62). ACM. https://doi.org/10.1145/2361276.2361288

Figueroa-Cañas, J., & Sancho-Vinuesa, T. (2019). Predicting early dropout students is a matter of checking completed quizzes: The case of an online statistics module. In M. Caeiro-Rodríguez, Á. Hernández-García, & P. J. Muñoz-Merino (Eds.), *Learning Analytics Summer Institute Spain 2019: Learning Analytics in Higher Education* (LASI 2019), 27–28 June 2019, Vigo, Spain (pp. 100–111). CEUR Workshop Proceedings. https://ceur-ws.org/Vol-2415/paper09.pdf

Foster, E., & Siddle, R. (2020). The effectiveness of learning analytics for identifying at-risk students in higher education. *Assessment & Evaluation in Higher Education*, *45*(6), 842–854. https://doi.org/10.1080/02602938.2019.1682118

Gardner, J., Yu, R., Nguyen, Q., Brooks, C., & Kizilcec, R. (2023). Cross-institutional transfer learning for educational models: Implications for model performance, fairness, and equity. In *Proceedings of the 2023 ACM Conference on Fairness, Accountability, and Transparency* (FAccT 2023), 12–15 June 2023, Chicago, Illinois, USA (pp. 1664–1684). ACM. https://doi.org/10.1145/3593013.3594107

Gašević, D., Dawson, S., Rogers, T., & Gasevic, D. (2016). Learning analytics should not promote one size fits all: The effects of instructional conditions in predicting academic success. *The Internet and Higher Education*, *28*, 68–84. https://doi.org/10.1016/j.iheduc.2015.10.002

Gupta, S. K., Antony, J., Lacher, F., & Douglas, J. (2020). Lean Six Sigma for reducing student dropouts in higher education—An exploratory study. *Total Quality Management & Business Excellence*, *31*(1-2), 178–193. https://doi.org/10.1080/14783363.2017.1422710

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, *77*(1), 81–112. https://doi.org/10.3102/003465430298487

He, J., Bailey, J., Rubinstein, B., & Zhang, R. (2015). Identifying at-risk students in massive open online courses. In *Proceedings of the AAAI Conference on Artificial Intelligence* (AAAI 2015), 25–30 June 2023, Austin, Texas, USA (Vol. 29). PKP|PS. https://doi.org/10.1609/aaai.v29i1.9471

Heron, M., Medland, E., Winstone, N., & Pitt, E. (2023). Developing the relational in teacher feedback literacy: Exploring feedback talk. *Assessment & Evaluation in Higher Education*, *48*(2), 172–185. https://doi.org/10.1080/02602938.2021.1932735

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. https://doi.org/10.1016/j.chb.2014.04.002

Jayaprakash, S. M., Moody, E. W., Lauría, E. J., Regan, J. R., & Baron, J. D. (2014). Early alert of academically at-risk students: An open source analytics initiative. *Journal of Learning Analytics*, *1*(1), 6–47. https://doi.org/10.18608/jla.2014.11.3

Jovanović, J., Saqr, M., Joksimović, S., & Gašević, D. (2021). Students matter the most in learning analytics: The effects of internal and instructional conditions in predicting academic success. *Computers & Education*, *172*, 104251. https://doi.org/10.1016/j.compedu.2021.104251

Kastberg, S. E., Lischka, A. E., & Hillman, S. L. (2020). Written feedback as a relational practice: Revealing mediating factors. *Studying Teacher Education*, *16*(3), 324–344. https://doi.org/10.1080/17425964.2020.1834152

Khalil, M., Slade, S., & Prinsloo, P. (2024). Learning analytics in support of inclusiveness and disabled students: A systematic review. *Journal of Computing in Higher Education*, *36*(1), 202–219. https://doi.org/10.1007/s12528-023-09363-4

Khan, I. A., Subhani, A., Rasheed, Z., Ahmad, U., & Brohi, M. N. (2023). Comprehensive assessment of risk assessment tools and academic performance in higher education: A meta-analytic perspective. *Journal of Applied Engineering & Technology (JAET)*, *7*(2), 10–24. https://doi.org/10.55447/jaet.07.02.116

Latif, A., Choudhary, A., & Hammayun, A. (2015). Economic effects of student dropouts: A comparative study. *Journal of Global Economics*, *3*(2), 1–4. https://www.hilarispublisher.com/open-access/economic-effects-of-student-dropouts-a-comparative-study-2375-4389-1000137.pdf

Liang, Z., Sha, L., Tsai, Y.-S., Gašević, D., & Chen, G. (2024). Towards the automated generation of readily applicable personalised feedback in education. In A. Olney, I. Chounta, Z. Liu, O. Santos, & I. Bittencourt (Eds.), *Artificial intelligence in education. AIED 2024. Lecture notes in computer science* (pp. 75–88, Vol. 14830). Springer. https://doi.org/10.1007/978-3-031-64299-9_6

Lin, J., Dai, W., Lim, L.-A., Tsai, Y.-S., Mello, R. F., Khosravi, H., Gasevic, D., & Chen, G. (2023). Learner-centred analytics of feedback content in higher education. In *Proceedings of the 13th International Conference on Learning Analytics and Knowledge* (LAK 2023), 13–17 March 2023, Arlington, Texas, USA (pp. 100–110). ACM. https://doi.org/10.1145/3576050.3576064

Liz Domínguez, M., Caeiro Rodríguez, M., Llamas Nistal, M., & Mikic Fonte, F. A. (2019). Predictors and early warning systems in higher education: A systematic literature review. In M. Caeiro-Rodríguez, Á. Hernández-García, & P. J. Muñoz-Merino (Eds.), *Learning Analytics Summer Institute Spain 2019: Learning Analytics in Higher Education* (LASI 2019), 27–28 June 2019, Vigo, Spain (pp. 84–99). CEUR Workshop Proceedings. https://ceur-ws.org/Vol-2415/paper08.pdf

López-García, A., Blasco-Blasco, O., Liern-García, M., & Parada-Rico, S. E. (2023). Early detection of students' failure using machine learning techniques. *Operations Research Perspectives*, *11*, 100292. https://doi.org/10.1016/j.orp.2023.100292

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. https://doi.org/10.3390/app10010354

Lu, O. H., Huang, J. C., Huang, A. Y., & Yang, S. J. (2017). Applying learning analytics for improving students engagement and learning outcomes in an MOOCs enabled collaborative programming course. *Interactive Learning Environments*, *25*(2), 220–234. https://doi.org/10.1080/10494820.2016.1278391

Lundberg, S. M., & Lee, S.-I. (2017). A unified approach to interpreting model predictions. *arXiv:1705.07874*. https://doi.org/10.48550/arXiv.1705.07874

Masabo, E., Nzabanita, J., Ngaruye, I., Ruranga, C., Nizeyimana, J. P., Uwonkunda, J., & Ndanguza, D. (2023). Early detection of students at risk of poor performance in Rwanda higher education using machine learning techniques. *International Journal of Information Technology*, *15*(6), 3201–3210. https://doi.org/10.1007/s41870-023-01334-3

Middleton, T., ahmed Shafi, A., Millican, R., & Templeton, S. (2023). Developing effective assessment feedback: Academic buoyancy and the relational dimensions of feedback. *Teaching in Higher Education*, *28*(1), 118–135. https://doi.org/10.1080/13562517.2020.1777397

Na, K. S., & Tasir, Z. (2017). Identifying at-risk students in online learning by analysing learning behaviour: A systematic review. In *2017 IEEE Conference on Big Data and Analytics* (ICBDA 2017), 16–17 November 2017, Kuching, Malaysia (pp. 118–123). IEEE. https://doi.org/10.1109/ICBDAA.2017.8284117

Nimy, E., Mosia, M., & Chibaya, C. (2023). Identifying at-risk students for early intervention—A probabilistic machine learning approach. *Applied Sciences*, *13*(6), 3869. https://doi.org/10.3390/app13063869

Osborne, J. B., & Lang, A. S. (2023). Predictive identification of at-risk students: Using learning management system data. *Journal of Postsecondary Student Success*, *2*(4), 108–126. https://doi.org/10.33009/fsop_jpss132082

Poulos, A., & Mahony, M. J. (2008). Effectiveness of feedback: The students' perspective. *Assessment & Evaluation in Higher Education*, *33*(2), 143–154. https://doi.org/10.1080/02602930601127869

Ryan, T., Henderson, M., Ryan, K., & Kennedy, G. (2021). Designing learner-centred text-based feedback: A rapid review and qualitative synthesis. *Assessment & Evaluation in Higher Education*, *46*(6), 894–912. https://doi.org/10.1080/02602938.2020.1828819

Schmidt, A., Cechinel, C., Queiroga, E. M., Primo, T., Ramos, V., Bordin, A. S., Mello, R. F., & Muñoz, R. (2025). Analyzing intervention strategies employed in response to automated academic-risk identification: A systematic review. *IEEE Revista Iberoamericana de Tecnologias del Aprendizaje*, *20*, 77–85. https://doi.org/10.1109/RITA.2025.3540161

Singell, L. D., & Waddell, G. R. (2010). Modeling retention at a large public university: Can at-risk students be identified early enough to treat? *Research in Higher Education*, *51*, 546–572. https://doi.org/10.1007/s11162-010-9170-7

Sonnleitner, B., Madou, T., Deceuninck, M., Theodosiou, F., & Sagaert, Y. R. (2025). Evaluation of early student performance prediction given concept drift. *Computers and Education: Artificial Intelligence*, *8*, 100369. https://doi.org/10.1016/j.caeai.2025.100369

Veerasamy, A. K., D'Souza, D., Apiola, M.-V., Laakso, M.-J., & Salakoski, T. (2020). Using early assessment performance as early warning signs to identify at-risk students in programming courses. In *Proceedings of the 2020 IEEE Frontiers in Education Conference* (FIE), 21–24 October 2020, Uppsala, Sweden (pp. 1–9). IEEE. https://doi.org/10.1109/FIE44824.2020.9274277

Waheed, H., Hassan, S.-U., Nawaz, R., Aljohani, N. R., Chen, G., & Gasevic, D. (2023). Early prediction of learners at risk in self-paced education: A neural network approach. *Expert Systems with Applications*, *213*, 118868. https://doi.org/10.1016/j.eswa.2022.118868

Wang, Q., & Mousavi, A. (2023). Which log variables significantly predict academic achievement? A systematic review and meta-analysis. *British Journal of Educational Technology*, *54*(1), 142–191. https://doi.org/10.1111/bjet.13282

Whitelock-Wainwright, A., Gašević, D., Tsai, Y.-S., Drachsler, H., Scheffel, M., Muñoz-Merino, P. J., Tammets, K., & Delgado Kloos, C. (2020). Assessing the validity of a learning analytics expectation instrument: A multinational study. *Journal of Computer Assisted Learning*, *36*(2), 209–240. https://doi.org/10.1111/jcal.12401

Winston, K. A., van der Vleuten, C. P., & Scherpbier, A. J. (2014). Prediction and prevention of failure: An early intervention to assist at-risk medical students. *Medical Teacher*, *36*(1), 25–31. https://doi.org/10.3109/0142159X.2013.836270

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. https://doi.org/10.1007/s40692-019-00143-7

Zacharis, N. Z. (2015). A multivariate approach to predicting student outcomes in web-enabled blended learning courses. *The Internet and Higher Education*, *27*, 44–53. https://doi.org/10.1016/j.iheduc.2015.05.002

Zhang, L., & Rangwala, H. (2018). Early identification of at-risk students using iterative logistic regression. In C. P. Rosé, R. Martínez-Maldonado, H. U. Hoppe, R. Luckin, M. Mavrikis, K. Porayska-Pomsta, B. McLaren, & B. du Boulay (Eds.), *Artificial intelligence in education. AIED 2018. Lecture notes in computer science* (pp. 613–626, Vol. 10947). Springer. https://doi.org/10.1007/978-3-319-93843-1_45

Zhang, Y., Fei, Q., Quddus, M., & Davis, C. (2014). An examination of the impact of early intervention on learning outcomes of at-risk students. *Research in Higher Education Journal*, *26*. https://files.eric.ed.gov/fulltext/EJ1055303.pdf