

RESEARCH ARTICLE

Open Access



An early warning system to identify and intervene online dropout learners

David Bañeres^{1*} , M. Elena Rodríguez-González¹, Ana-Elena Guerrero-Roldán¹ and Pau Cortadas²

*Correspondence:
dbaneres@uoc.edu

¹ Faculty of Computer
Science, Multimedia
and Telecommunications
Department, UOC, Barcelona,
Spain

² Faculty of Economics
and Business, UOC, Barcelona,
Spain

Abstract

Dropout is one of the major problems online higher education faces. Early identification of the dropout risk level and an intervention mechanism to revert the potential risk have been proved as the key answers to solving the challenge. Predictive modeling has been extensively studied on course dropout. However, intervention practices are scarce, sometimes mixed with mechanisms focused on course failure, and commonly focused on limited interventions driven mainly by teachers' experience. This work contributes with a novel approach for identifying course dropout based on a dynamic time interval and intervening, focusing on avoiding dropout at the assessable activity level. Moreover, the system can recommend the best interval for a course and assessable activity based on artificial intelligence techniques to help teachers in this challenging task. The system has been tested on a fully online first-year course with 581 participants from 957 enrolled learners of different degrees from the Faculty of Economics and Business at the Universitat Oberta de Catalunya. Results confirm that interventions aimed at goal setting on the ongoing assessable activity significantly reduce dropout issues and increase engagement within the course. Additionally, the work explores the differences between identification mechanisms for course dropout and failure aiming to distinguish them as different problems that learners may face.

Keywords: Online learning, Early warning system, Dropout, Intervention, Artificial intelligence

Introduction

Nowadays, teachers are pushed to provide a better personalized educational experience based on learners' needs. Some examples of such personalization are feedback on assessable activities, resources recommendation, and self-regulating guidelines (El-Sabagh, 2021). However, these strategies are more effective when they are provided based on learners' individual needs, such as at-risk situations that may appear. Although there are many possible at-risk situations, we can summarize them into two possible ones: at-risk of failing or dropping out of the course (Simpson et al., 1980).

Dropout is a challenging problem in Higher Education (HE), including face-to-face, blended and online settings (Grau-Valldosera & Minguillón, 2014; Stone & O'Shea, 2019; Yair et al., 2020), since it seems an unsolvable problem. Many researchers have analyzed the problem (Greenland & Moore, 2022; Xavier & Meneses, 2020), and the factors

(Bağrıacık Yılmaz & Karataş, 2022; Greenland & Moore, 2022; Thalhammer et al., 2022; Tinto, 1975; Xavier & Meneses, 2022) to enlighten the reasons behind the problem, but learners are still dropping out.

The first problem is that there is no agreement on the dropout definition (Xavier & Meneses, 2020) and sometimes misinterpreted with related concepts (e.g., completion, retention, success, persistence, among others). Moreover, dropout has a temporal conception ranging from a long-term perspective (leaving an academic program or university) to a mid-term perspective (course dropout). The second problem is the vast list of factors that may influence dropout. According to (Lee & Choi, 2011) dropout factors can be classified into learner-related, course-related, and external factors. Learner-related focuses on prior knowledge (Park & Choi, 2009) and learners' traits such as determination (Eriksson et al., 2017), self-efficacy (Hart, 2012), time pressure (Xavier & Meneses, 2022), motivation (Douglas et al., 2020) or self-regulation (de Barba et al., 2020), but there is no consensus on how demographic characteristics affect dropout (Kizilcec et al., 2016). Course-related factors involve, for instance, support deficit from teachers (Bakar et al., 2020; Hone & el Said, 2016; Ross & McNealy, 2020) or unaccomplished course expectations (Darkenwald & Gavin, 1987). External factors are unexpected, such as family and work commitments (Eriksson et al., 2017).

Course-related factors can be detected through interviews and learners' opinion surveys at the end of the course. It may require longitudinal analyses to seek better teaching strategies or learning resources. Learner-related factors are distinct for each learner but also detectable and the ones which can be impacted to avoid dropout issues. For example, identification supported by learner's profile analysis provides an initial approach to detecting prior-knowledge shortcomings (NeCamp et al., 2019). However, at-risk identification based on artificial intelligence (AI) techniques, including statistical models, educational data mining, and machine learning, has a major potential to identify other factors (Zawacki-Richter et al., 2019) such as failure or dropout.

This work focuses on dropout identification from a mid-term perspective within a specific course in an online HE setting. We define a dropout learner as "a learner who does not submit the ongoing assessable activity" because not submitting such activities is a high indicator of dropping out of a course (Rodríguez et al., 2019). Learner-related factors are assessed based on the learner's profile, engagement, and performance within the course. In the case of potential dropout identification, an intervention mechanism is triggered to impact the learner's traits, such as motivation and self-regulation.

The work presented in (Guerrero-Roldán et al., 2021) discussed an Early Warning System (EWS) combined with an intervention mechanism to revert course failure in at-risk situations. The system, denoted as Learning Intelligent System (LIS), provided the first step to support learners within a course by providing early identification of at-risk learners and sending personalized messages when needed as part of an intervention mechanism (Rodríguez et al., 2022). A positive effect was shown on learners who used the system. However, the system lacked a procedure to early distinguish dropout learners who were identified too late, when the dropout was already materialized.

The contribution of this paper is twofold. First, the LIS system is enhanced by the early identification of course dropout learners. A dynamic time interval in days is proposed for each assessable activity, and a daily prediction is performed for each learner. A

learner is identified as a potential dropout when she is predicted as a dropout during a consecutive number of days specified by the time interval. New dashboards are provided to teachers and learners to increase awareness about the potential dropout risk based on a new predictive model. Second, the intervention mechanism has been enhanced with a specific intervention when a learner's likelihood of dropping out is signaled.

The paper is organized as follows. The next section provides the theoretical framework for dropout identification and intervention mechanisms, as well as the background of the university educational model. The third section presents the methodological approach, the study procedure, the participants, the used data, and the instruments. The fourth section describes the obtained results and the main findings. Finally, the last section provides conclusions, limitations, and future research lines.

Theoretical framework and background

Dropout identification and intervention mechanisms

Dropout identification in online settings has been extensively studied in Massive Open Online Course (MOOC) settings (Dalipi et al., 2018; Goel & Goyal, 2020; Moreno-Marcos et al., 2019) due to the low retention. Models have high effectiveness in identifying dropout learners (Mubarak et al., 2020; Whitehill et al., 2017) since models are trained with unbalanced data where dropout distribution ranges between 70 and 90% on average. Some of these models have been used to develop analytical tools to help teachers to understand when dropout appears and which factors related to engagement and performance impact its materialization (Boudjehem & Lafifi, 2021; Chen et al., 2017; Dourado et al., 2021; Itani et al., 2018; Tang et al., 2015). Analytical tools increase awareness but have a low influence on reverting the at-risk situations.

Not all dropout factors can be intervened. External factors such as family and work commitments or unexpected events may impact dropout and are unavoidable. However, course-related factors are associated with learners' expectations and perceptions. Learners may feel unattended (Hart, 2012; Hone & el Said, 2016), impersonal supported (Henry, 2018), overwhelmed by perceived course difficulty (Greenland & Moore, 2022; Xenos et al., 2002) or without time management skills (Veletsianos et al., 2021). Those factors can be influenced by improving learner-related factors. Motivation (Borrella et al., 2019; Douglas et al., 2020), self-regulation (Stephen et al., 2020), and self-efficacy (Bandura, 1997) have been found as relevant factors to focus the interventions. Since learners' factors are highly connected, improving motivation, self-regulation, and self-efficacy increase persistence (Xavier & Meneses, 2022) which has been found as a personal trait that mitigates dropout. Persistence affected by self-determination to complete a specific goal (i.e., knowledge growth, career advancement, leisure) is a relevant factor to success (Broadbent & Poon, 2015).

Interventions seem to answer to the dropout problem (Xavier & Meneses, 2022). Different approaches have been explored in MOOC settings, and automated interventions have been sought due to the minimal available teaching support in such environments. Authors in (Borrella et al., 2022) proposed an intervention based on adapting the difficulty of the assessable activities for learners who stopped submitting quizzes to impact motivation and self-efficacy. An automated system to propose additional resources was developed for a language programming course to promote self-efficacy (Teusner et al.,

2018). However, effective communication with the teacher has been found as one of the best positive strategies for course retention (Hart, 2012). According to (Thaler & Sunstein, 2008), nudges are an effective way to influence individuals' behavior and decisions, and different intervention systems have been developed using such an approach. Authors in (Kurtz et al., 2022) proposed a nudging intervention system when a learner did not submit a quiz for helping to improve self-regulation and self-efficacy, and just-in-time interventions were proposed in (Teusner et al., 2018) focusing on motivating learners. Interventions do not always positively impact learners (Borrella et al., 2019). However, the negative impact in this last study case was influenced by the interval when the intervention was sent (e.g., only one intervention on a mid-term exam) instead of the message content.

Consequently, timely interventions are crucial to impact potential dropout learners. Using the wrong interval may cause confusion and demoralization (Woodley & Simpson, 2013), negatively affecting dropout factors. For instance, authors in (Boudjehem & Laffi, 2021) proposed message interventions when a particular learning objective was not accomplished in the middle of the course focusing on self-regulation. Fixed temporal intervals have also been proposed, such as week intervals (NeCamp et al., 2019) or a percentage of the course duration (Borrella et al., 2019), targeting both approaches to motivate learners, mostly adjusted to the assessable activities' duration in MOOC settings. However, all experiences considered the teachers' experience to decide the best interval to perform the interventions.

The aim of the interventions is also relevant. Several types of feedback can be provided and are not mutually exclusive: cognitive, behavioral, outcome-oriented, and process-oriented (Sedrakyan et al., 2020). Note that feedback may affect different factors. Cognitive and outcome-oriented (e.g., How do I perform?) improve knowledge and, therefore, impact confidence and motivation. Behavioral and process-oriented (e.g., How can I do better?) target self-efficacy and self-regulation, and, in the end, persistence to continue the course. Cognitive, outcome and process-oriented feedback can only be given when an activity has been assessed. Some methodologies are based on such types of feedback (Kurtz et al., 2022; Teusner et al., 2018). However, the intervention arrives too late when the dropout is materialized. Early behavioral feedback combined with goal setting (Latham & Locke, 2007; Locke & Latham, 2002) or self-goal setting (Elliot & Fryer, 2008; Zimmerman, 1990) has been proved to be a dropout mitigation strategy by promoting motivation, self-efficacy, and self-regulation (Jivet et al., 2020).

Previous works mainly focused on MOOC environments. Although the intervention messages (and the attached content) can be widely applied in different settings, parameters such as the predictive model for early identification and the intervention interval are not fully applicable in online HE settings. Predictive models are affected mainly by the used data, and their behavior will differ depending on their applied settings. Interval for interventions is also different. The duration of assessable activities in MOOC settings usually is a week, and interventions are adjusted to such period. However, in non-MOOC settings, the assessable activities duration is commonly variable, which may significantly affect when the intervention should be made. Authors in (Burgos et al., 2018) simplified the problem by sending messages and performing telephone call interventions on 20%, 35%, and 50% of the course period. Another intervention approach was

tested in (Figuroa-Cañas & Sancho-Vinuesa, 2021) where a unique intervention was done at 25% of the course period based on messaging learners and extending the submission date of previous quizzes.

Our work improves previous approaches by training a predictive model specifically for each course and providing a Dynamic Temporal Window (denoted as DTWin) for each assessable activity. The DTWin is an interval of consecutive days. A learner is considered a dropout in the assessable activity when she is predicted as a dropout during a consecutive number of days specified by the DTWin size. Moreover, the EWS recommends the optimal DTWin size for a course and assessable activity based on the predictive model. Thus, an intervention can be triggered for each assessable activity when a learner is considered a potential dropout before the submission date of the ongoing assessable activity. The EWS transforms the prediction into an explainable risk-level identification status for the teachers and learners in the dashboards. The intervention mechanism, fully customizable by the teachers, can trigger a personalized message based on the teachers' experience. Although the intervention mechanism provides templates for the messages, teachers are encouraged to enrich the messages with their knowledge and experience about the resources, exercises, and shortcoming goals the learner should complete to try to revert possible dropout at-risk situations targeting motivation, self-efficacy, and self-regulation by behavioral feedback.

A fully online university

The Universitat Oberta de Catalunya (UOC) is a fully online university born in 1994. The educational model is centered on the learner focusing on the competencies to be acquired. The assessment process is based on a continuous assessment (CA) model combined with a summative assessment at the end of the semester. Therefore, the CA includes different Continuous Assessable Activities (CAA) during the semester and a final examination at the end of the course. The course final mark is computed based on a predefined formula where each CAA has a different weight depending on its significance within the course. The grading system is based on qualitative scores on CAA. Each assessable activity is graded with the following scale: A (very high), B (high), C+ (sufficient), C− (low), and D (very low), where a grade of C− and D means failing the CAA. In addition, another grade (N, non-submitted) is used when a student does not submit the CAA. A number between 0 and 10 scores the CA and the final mark, where a grade lower than five means failing the CA and the course, respectively.

Teachers and learners communicate through the VLE, which includes different communication spaces to promote learning and social interactions. The teaching is performed in online classrooms that include the learning materials, tools, CAA, and communication spaces (i.e., teacher's blackboard, forum, and debate). A relevant remark about the VLE is that all learners' traces within the VLE are stored within an institutional data mart (Minguillón et al., 2018) including learners' historical and current semester data anonymously. The system also stores learners' online behavior in terms of submitted CAA, performance, and clickstream (i.e., navigational data, accessed resources, and tools utilization), among others.

Learners mostly have a full-time job and familiar commitments, so they have some time constraints to overcome when enrolling in online courses. They pursue new studies

to expand knowledge related to a specific domain or improve their professional careers. Teachers design CAA and provide feedback to learners to reduce isolation and guide learners as much as possible across their learning path. The LIS system is currently being developed to improve learners' awareness about their at-risk status, reinforce self-efficacy and motivation by setting shortcoming goal settings and, eventually, enhance personalization. In this work, the baseline LIS system, with a course failure identification mechanism (Baneres et al., 2020) combined with an intervention mechanism (Rodríguez et al., 2022), has been improved with a new dropout risk-level identification and a dropout intervention mechanism. The contribution of this paper has been evaluated by answering the following research questions when testing the system on a specific first-year HE course:

RQ1. How accurate is the dropout risk-level identification on the LIS system?

RQ2. Is dropout decreased when using the LIS system?

RQ3. Is there any relationship between dropout and failure risk-level identification mechanisms?

Methodology

Research design and participants

The research methodology behind the development of the LIS system uses a mixed research design. An action research methodology (Oates, 2006) combined with a design and creation approach (Kuechler & Vaishnavi, 2012) has been used. The former is used because a product needs to be developed and tested in real-case learning scenarios. In this case, the product is a system that aims to solve a problem in teaching–learning environments. The principles of an action research methodology are to focus on practical issues, an interactive plan-act-reflect cycle, collaboration with practitioners, and data generation methods to evaluate the product outcomes. The latter involves the creation of new Information Technology (IT) artifacts. The methodology is an iterative problem-solving approach composed of five steps (Kuechler & Vaishnavi, 2012): awareness of the problem to be solved; suggestion of tentative ideas to solve it, development of an IT artifact based on the suggested ideas, evaluation of the artifact whether it meets the expectations, and conclusions about the results to gain knowledge.

This paper focuses on the second iteration of developing the dropout intervention mechanism integrated within the LIS system. In the first iteration, the dropout predictive model was designed and compared with previous approaches proposed by other authors. The second iteration involves the integration of the model into the teaching–learning process by designing an intervention mechanism for dropout at-risk learners. In this cycle, the dropout intervention mechanism has been tested in an online course of 6 ECTS called *Markets and behavior* of the Faculty of Economics and Business.

Markets and behavior is an introductory course included in the specialty of micro-economy and mandatory within the Faculty of Economics and Business. This course facilitates the comprehension of the characteristics and adjustment measures in the modern economy based on the interaction between supply (companies and their costs) and demand (consumers and their preferences). The assessment model comprises five

CAA combined with a Validation Test (VT) for learners who passed the CAA. The condition to be eligible to perform the VT is to pass the CA with a score greater than five and submit four out of the five CAA. The final mark FM of the course is computed as $FM = 70\% CA + 30\% VT$. If the CA is not passed or the learner decides not to do the course through the CA, the learner can perform a final exam where FM is the mark of the exam. When performing the CAA, a learner receives individual feedback when she fails an activity or upon request. This course has been selected because it is a first-year mandatory course within different faculty degrees (i.e., BSc. in Economics, BSc. in Business Management and Administration, BSc. in Tourism, and BSc. in Market Research and Marketing) with a large number of enrolled learners. The course also offers some flexibilization strategy (i.e., four out of the five CAA have to be submitted) that could mitigate even more dropout issues combined with the proposed mechanism. Finally, reducing dropouts in this course can also be advantageous at the degree level. However, this last assumption is out of the scope of this paper and requires a longitudinal analysis of cohorts after several semesters.

The research design is illustrated in Fig. 1. The dropout intervention mechanism has been tested on the *Markets and behavior* course during the 2021 Spring semester (i.e., from February to June 2021). Learners interested in participating signed a consent form to participate in the pilot because the institutional Research Ethical Committee requires an explicit acceptance to be included in any study, following the General Data Protection Regulation (GDPR, <https://gdpr-info.eu/>). The consent form informs about the data the system collects and processes and its capabilities. After acceptance, learners will receive a daily prediction about the likelihood of dropping out (i.e., not submitting the ongoing CAA) in a personalized dashboard. If a learner is identified early as at

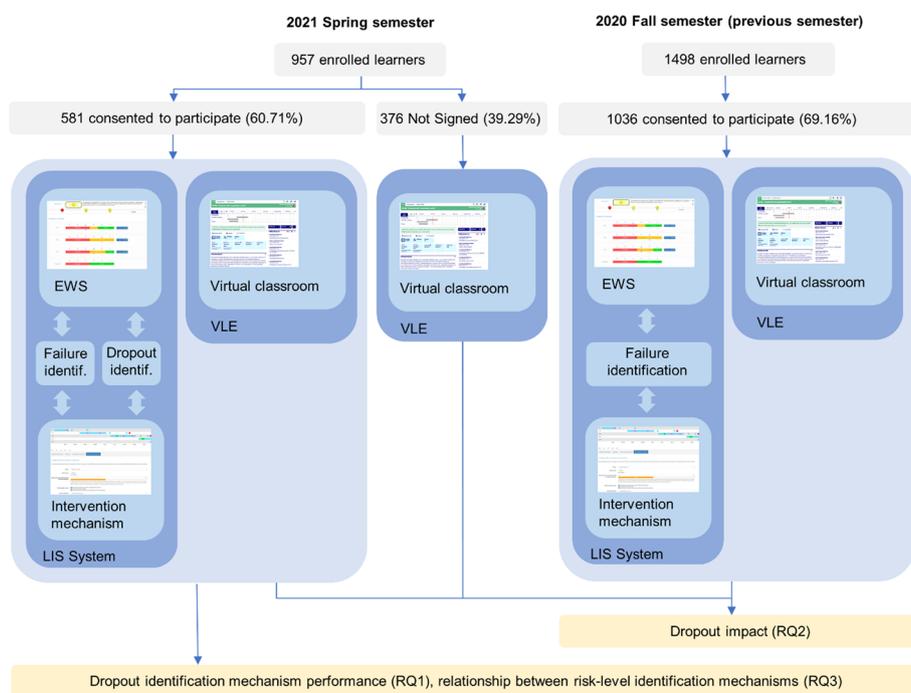


Fig. 1 Research design and participants

risk, the intervention mechanism will automatically provide the learner with meaningful information to help submit the ongoing CAA. Figure 1 shows the high participation of 60.71% of the learners (i.e., 581 from 957). Note that the LIS system also incorporates a course failure identification mechanism, whose impact was analyzed in (Guerrero-Roldán et al., 2021), and it is the baseline system used in the previous semester (i.e., 2020 Fall semester). We gathered the learners' performance results from the previous semester who consented to use the course failure identification mechanism. The participants' conditions were the same, but learners only received information from the intervention mechanism to avoid course failure. There was also high participation of 69.16% (i.e., 1036 from 1498). After finishing the course, the performance of the dropout risk-level identification mechanism is analyzed (i.e., the percentage of correct dropout identifications regarding the submission event) to answer the RQ1; the impact of the intervention mechanism is evaluated by comparing dropouts on participants versus learners who do not consent and learners from previous semester who only tested the failure identification mechanism to answer RQ2; and the correlation between both risk-level identification mechanisms to answer RQ3.

Study procedure and instruments

As previously described, this iteration involves integrating the predictive approach to detect potential dropout learners with the intervention mechanism within the LIS system.

The predictive approach involves a two-step method. In the first step, a predictive model denoted as PDAR (Profiled Dropout At Risk) is trained for each day of the course. Such a model takes into account learner's profile information (i.e., number of enrolled courses in the current semester, number of repeated enrollments in the target course, whether she is a novice learner, and the grade point average of the academic report), performance data within the course (i.e., the score of the already assessed CAA, and whether previous CAA have been submitted), and daily clickstream data about the VLE utilization (i.e., access to the VLE, the classroom of the target course, the resources, and tools; and number of messages read in the forum and teacher's blackboard). The model's outcome is the likelihood of not submitting the ongoing CAA. Aforesaid, we defined the learner's dropout risk as the risk of not submitting the ongoing CAA. Not submitting an activity in a CA model is evidence that the learner may drop out eventually from the course. The outcome is a binary variable: 1—not submitting, 0—submitting. However, such a model suffers a low accuracy in detecting non-dropout learners because learners typically do not access the VLE daily.

Therefore, DTWin is proposed for each CAA which improves the accuracy of identification by reducing the false positive cases (i.e., non-dropout learners incorrectly identified as dropouts). Recall that the DTWin is an interval of consecutive days defined for each CAA. A learner is considered a dropout in a CAA when she is predicted as a dropout by the PDAR model during a consecutive number of days larger than the specified DTWin size. Although teachers can manually define the DTWin size based on their preferences and experience, the LIS system can provide a recommended DTWin size based on an optimization procedure that maximizes the correct identification of dropout and non-dropout learners. The optimization process searches in the testing dataset

(i.e., learners of the previous semester) for each learner the largest interval of consecutive days that the learner is predicted by the PDAR model as a dropout and gathers whether she has not submitted the CAA. Then, the process identifies for each possible DTWin which learners have been predicted as dropouts by varying the window size from one day to the total number of days of the CAA. When the explored window size fits in the largest dropout interval for a learner (i.e., the largest interval of consecutive days predicted as a dropout is larger than the window size), the learner’s prediction is set to the likelihood of dropping out for such a window. Finally, TPR and TNR’s metrics can be computed from the predicted and actual dropout events for each window size. The best window size (i.e., the DTWIN size) is selected by maximizing the sum of the TPR (i.e., accuracy when detecting at-risk dropout learners) and TNR (i.e., accuracy when detecting non-at-risk ones).

Based on the PDAR predictive model, the LIS system computes a dropout prediction for each learner and day of the course. This information is transformed into a risk level based on different colors based on the decision tree shown in Fig. 2. This decision tree has been built based on experimentation on the failure identification system. The decision tree applies to all university courses and differentiates between certainly and possible predicted likelihood dropout events depending on the DTWin accuracy on the target course. Note that course personalization is achieved by the course’s corresponding dropout model and the teacher’s intervention messages. The risk level is set to an intermediate risk level when the accuracy is low due to possible misprediction. Additionally, the decision tree considers the non-submission event of the previous CAA as a relevant event. Even though a learner has not submitted one CAA, she can still be active and receive interventions. However, when the learner does not submit more than one CAA, the learner has probably dropped out of the course, and no intervention is triggered. From previous pilots, we observed that there is room for improvement when a learner does not submit one CAA but not when more than one (Rodríguez et al., 2022). The colors differ for learners and teachers. While learners only have a three-color risk level following a traffic light metaphor, teachers have more colors to know each learner’s

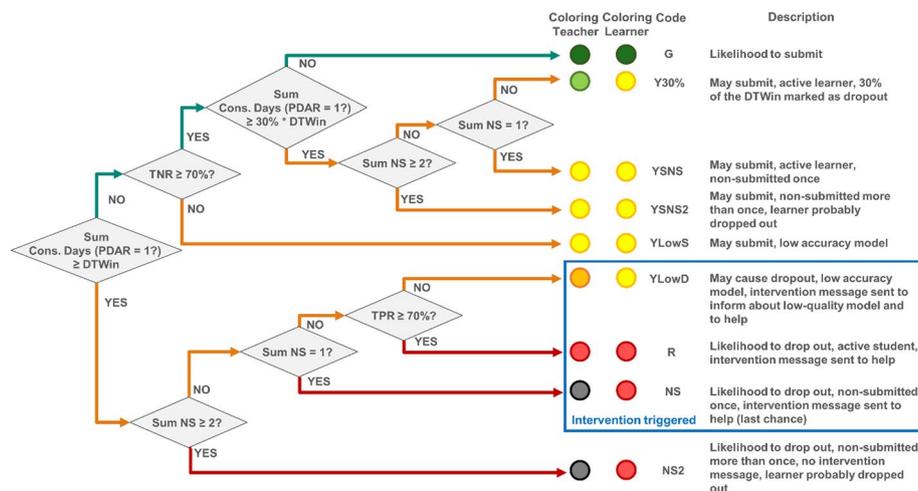


Fig. 2 Decision tree for the dropout risk-level classification

status better. The green color is given to low-risk learners and learners who do not reach 30% of the DTWin size. After different experiments, we observed that most active learners tended to access VLE below such a limit. This decision tree has been built based on experimentation on the failure identification system. The decision tree applies to all university courses and differentiates between certainly and possible predicted likelihood dropout events depending on the DTWin accuracy on the target course. Note that course personalization is achieved by the course’s corresponding dropout model and the teacher’s intervention messages. The risk level is set to an intermediate risk level when the accuracy is low due to possible misprediction. Additionally, the decision tree considers the non-submission event of the previous CAA as a relevant event. Even though a learner has not submitted one CAA, she can still be active and receive interventions. However, when the learner does not submit more than one CAA, the learner has probably dropped out of the course, and no intervention is triggered. From previous pilots, we observed that there is room for improvement when a learner does not submit one CAA but not when more than one (Rodríguez et al., 2022). The yellow color for learners represents a medium risk with different meanings. This color is set for learners transitioning to the DTWin size limit where not submitting some previous CAA is also considered (i.e., Y30%, YSNS, and YSNS2 levels). It is also assigned to activities with low-quality predictive models (i.e., where the accuracy of detecting at-risk learners—TPR and non-at-risk learners—TNR are lower than a threshold of 70%). A low-quality model cannot guarantee the prediction’s correctness; therefore, high and low risks are set to medium. Finally, the red color represents a high-risk level when the DTWin size is reached, and the predictive model is a high-quality one (i.e., the TPR larger than 70%). Note that low-quality models for detecting dropouts have an orange color for teachers to distinguish at-risk levels. Red color (and black for teachers) is also assigned to at-risk learners who did not submit previous CAA (i.e., represented in the decision nodes as not submitted—NS and NS2).

This information is provided to teachers and learners in dashboards. Figure 3 illustrates the teacher’s dashboard where the two risk level mechanisms are provided: course failure and dropout. For each CAA, the risk level of course failure (Baneres et al., 2020)

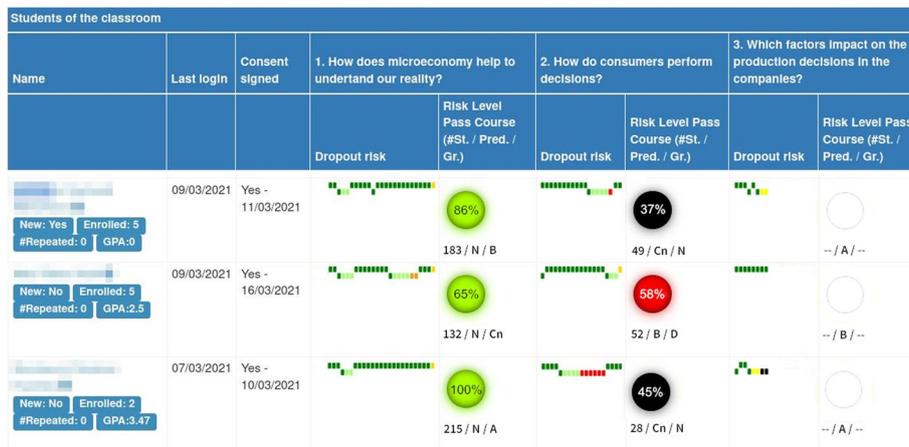


Fig. 3 Teacher’s dashboard during the CAA3

is shown in a green-amber-red-black color distribution with the percentage of learners passing the course based on historical data. The number of learners in the historical data with a similar risk level, the predicted grade to have a medium or low risk of failing, and the grade finally obtained by the learner are also shown. The course failure risk level for the ongoing CAA is computed when the previous CAA is assessed. For example, we can observe in the figure that the risk level for the CAA3 has been computed. This information is ready while the learners are performing the CAA3 and when the teacher has assessed the CAA2. Thus, the learner knows before submitting the CAA which grade must obtain to likely pass the course (i.e., medium- or low-risk level). This goal-setting information intends to motivate the learner to reach the minimum suggested grade.

The dropout risk level is based on the dropout classification tree shown in Fig. 2. A dropout risk level is provided for each day of the ongoing CAA showing an evolution through a sparkline chart. Each prediction has contextual help to inform the teacher about the detailed description of the risk. Figure 3 shows many of the risk levels. For instance, taking into account that the CAA1 has a low-quality model for *Markets and behavior* (i.e., TPR smaller than 70%), the YLowD risk level (i.e., orange color) is delivered to the second learner instead of the R high-risk level when the DTWin size is reached. Other CAA have a high-quality model that allows assigning the R high-risk level (i.e., red color) to the first and third learner in the CAA2. Those learners did not submit the CAA2. Therefore, we can observe that YSNS and NS risk levels (i.e., yellow and black colors, respectively) are assigned in the CAA3 in the current day. Note that when the learner submits the CAA, the sparkline shows the event with golden color, and the dropout risk level computation is deactivated for such learner until the next CAA.

A learner also receives information about her potential risk levels. Figure 4 illustrates the main part of the dashboard for the first learner of Fig. 3, where the risk-level semaphores for each risk level are shown together with information about personal

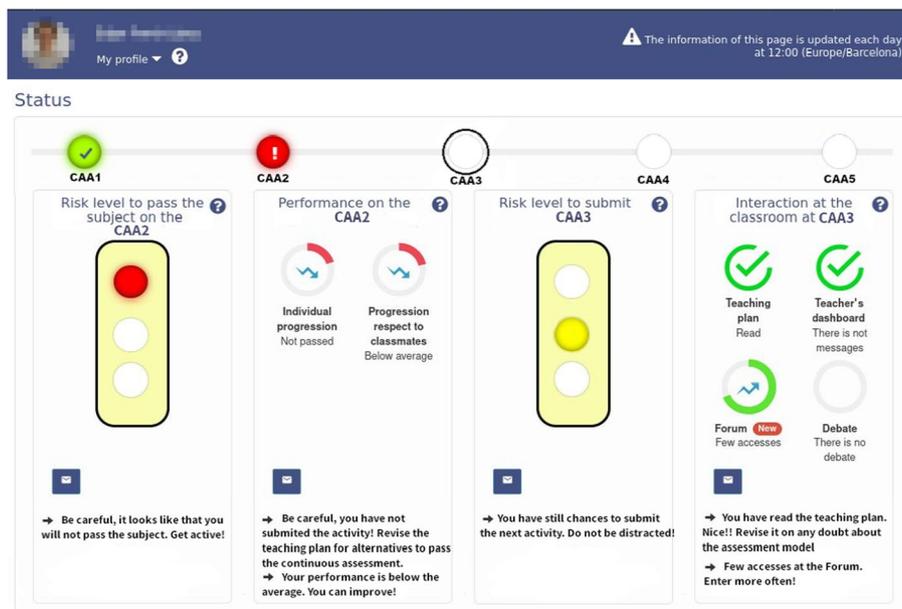


Fig. 4 Learner's dashboard during the CAA3

progression and comparison with other course learners. Each piece of information also provides additional explanations to inform the learner about her status.

The dashboard also provides detailed information about risk-level identification. Figure 5 illustrates such part. The *risk level for activities* section summarizes the information for the risk of course failure. In assessed CAA, the learner observes the risk level where her grade is, and for the ongoing CAA, the grades distribution for each risk level. The dropout risk is shown in the *risk level of not submitting the current activity* section. Here, the risk is drawn on an at-risk bar divided by as many points as the DTWin size has. The DTWin size and, therefore, the number of specific points in the bar are not specified to the learner to avoid technical details. The learner only observes the proximity of the at-risk level marked by two triangles which, in fact, indicate the number of consecutive days the learner has been predicted as a dropout for the corresponding PDAR model of the CAA. In the example, the DTWin size of the CAA3 is four days, and the learner has been predicted as a dropout for three consecutive days. If the PDAR predicts her as a dropout on the following day, the DTWin size will be reached, and therefore, the learner will be considered a potential dropout, and the red color will be set in the dropout risk semaphore.

Finally, the system has a built-in intervention mechanism that only sends messages to learners on YLowD, R, and NS risk levels, as shown in Fig. 2. As previously discussed, such risk levels are the ones where the messages may have some impact on reverting the potential dropout issue. The messages are triggered automatically by the system when such risk levels are detected on behalf of the teacher. However, the content of the messages has been previously designed and written by the teachers because their expertise helps to enhance the goal-setting objectives of the messages. In order to increase personalization, there is a different message for each risk level and CAA.



Fig. 5 Detailed risk-level information for the learner's dashboard during the CAA3

The LIS system has been tested in the *Markets and behavior* course during the 2021 Spring semester. The learners receive a daily prediction about their risk of not submitting the next CAA jointly with the prediction of passing the course after assessing each CAA. In case of being at risk, the intervention mechanism sends the corresponding message to the learner. This paper analyses the dropout risk-level identification and intervention mechanism by answering the research questions stated before.

Data analysis

We used data from different sources to answer the research questions that arose in this paper. First, records from the LIS system have been used to analyze the performance of the DTWin approach (RQ1) and the relationship with the risk level of course failure (RQ3). Also, such data have been combined with anonymized data from the institutional data mart for analyzing the engagement of participants after receiving an intervention message (RQ1) and the dropout impact (RQ2).

The performance of the DTWin approach uses the following metrics:

$$TNR = \frac{TN}{TN + FP} \quad ACC = \frac{TP + TN}{TP + FP + TN + FN}$$

$$TPR = \frac{TP}{TP + FN} \quad F_{1.5} = \frac{(1 + 1.5^2)TP}{(1 + 1.5^2)TP + 1.5^2FN + FP}$$

where TP denotes the number of at-risk students correctly identified, TN the number of non-at-risk students correctly identified, FP the number of non-at-risk students not correctly identified, and FN the number of at-risk students not correctly identified. These four metrics are used for evaluating the global accuracy of the model (ACC), the accuracy when detecting at-risk learners (true positive rate—TPR), the accuracy when distinguishing non-at-risk learners (true negative rate—TNR), and a harmonic mean of the true positive value and the TPR that weights correct at-risk identification (F score—F1.5). The accuracy of the dropout risk-level identification is computed by checking the percentage of correct dropout identifications regarding the learners who finally have not performed the VT or final exam.

Statistical analysis has also been used to answer the research questions. Results have been computed with R language scripts embedded in the LIS system. Boschloo's unconditional test has been performed to see the association between submitting the CAA and participating in the pilot represented as binary variables (RQ2). Note that no association would indicate that the submitting event is not conditioned by participating in the pilot and, therefore, using the intervention mechanism. Moreover, a Chi-squared test is performed to check the independence between both risk-level identification mechanisms (RQ3). In the case of dependency, the Chi-squared test does not give enough evidence about the correlation among variables (i.e., risk levels identification mechanisms) with more than two levels. Therefore, the Cramer's V test is used to measure the percentual correlation between both risk-level identification mechanisms.

Results

RQ1. How accurate is the dropout risk-level identification on the LIS system?

Before analyzing the accuracy of the dropout risk-level identification mechanism, Table 1 summarizes the performance of the best DTWin size selection for each CAA. The model has been trained with data from the 2017 Fall to 2020 Spring semesters and validated with data from the 2020 Fall semester. The table shows the *activity duration*, the selected window for detecting potential dropout learners (*Best DTWin*), and the performance metrics for each CAA. The selected window is the recommended one for the LIS system and ranges from 25 to 33% of the CAA duration. Such intervals produce high-quality identification models, except for the CAA1 and CAA5 for detecting dropout learners (i.e., such models will trigger YLowD risk level instead of R regarding the decision tree of Fig. 2).

When analyzing the identification of dropout learners, the TPR has a considerable high value larger than 65% from the CAA1 and reaches a value larger than 85%, except in the last activity, CAA5. We can observe that it is easier to detect non-at-risk learners (i.e., TNR) since such learners tend to regularly access the online classroom and communication spaces and pass CAA activities. However, we can see that, surprisingly, the performance model for the last CAA5 is worse than the others. The reason raised in (Guerrero-Roldán et al., 2021) is due to the flexibilization strategy in the assessment model. Recall that learners can pass the CA by submitting four out of the five CAA. Thus, we can distinguish three types of learners in the last CAA5. First, outstanding learners submit all five activities to get the maximum score. Second, some learners have not submitted some previous CAA and submit the last one to be eligible to pass the course with the VT. Finally, some learners have submitted the four previous activities and decide not to submit the last one. The results show that it is difficult for a model to predict the behavior of the different types of learners.

Table 2 summarizes the model’s performance on the semester pilot on consented learners (i.e., 2021 Spring semester). The TNR metric behaves similarly to the validation test when detecting non-at-risk learners. However, the TPR performance has decreased significantly.

In order to get insights into the reason why the TPR performance of the DTWin approach is substantially low, Table 3 provides information about the average access time per day for each group of identified learners. The table summarizes the number of learners (*n*.) and the average access time to the online classroom. The average time is shown for the complete activity (during activity) for the detected learners as non-at-risk (i.e., FN and TN groups). It is split into two average times (before and after) for

Table 1 Performance of the best DTWin on the validation test

Activity	Activity duration	Best DTWin	TN	FN	TP	FP	ACC (%)	TNR (%)	TPR (%)	F _{1,5} (%)
CAA1	24	6	671	60	122	237	72.75	73.90	67.03	51.59
CAA2	18	6	729	79	191	91	84.40	88.90	70.74	69.78
CAA3	14	4	831	34	207	18	95.22	97.87	85.89	87.68
CAA4	16	5	785	59	239	7	93.94	99.11	80.20	84.75
CAA5	15	4	289	289	405	107	63.67	72.97	58.36	63.47

Number of learners (n = 1090)

Table 2 Performance of the best DTWin on the semester pilot

Activity	TN	FN	TP	FP	ACC (%)	TNR (%)	TPR (%)	F _{1.5} (%)
CAA1	378	15	16	172	67.81	68.73	51.61	20.17
CAA2	381	73	46	81	73.49	82.47	38.66	37.87
CAA3	459	34	64	24	90.02	95.03	65.31	67.42
CAA4	446	46	81	8	90.71	98.24	63.78	70.25
CAA5	177	217	139	48	54.39	78.67	39.04	45.72

Number of learners who consented (n = 581)

Table 3 Average access to the online classroom during activity duration by type of identified learner

Activity		Non-at-risk incorrectly identified (FN)	Non-at-risk correctly identified (TN)		At-risk incorrectly identified (FP)	At-risk correctly identified (TP)
CAA1	n	15	378	n	172	16
	Before			0.09		0.06
	During activity	0.31	0.36	After	0.23	0.16
CAA2	n	73	381	n	81	46
	Before			0.15		0.20
	During activity	0.27	0.36	After	0.24	0.11
CAA3	n	34	459	n	24	64
	Before			0.16		0.09
	During activity	0.26	0.40	After	0.20	0.05
CAA4	n	46	446	n	8	81
	Before			0.09		0.07
	During activity	0.24	0.40	After	0.27	0.03
CAA5	n	217	177	n	48	139
	Before			0.40		0.13
	During activity	0.40	0.61	After	0.41	0.07

Average access is the average access time by learner and day

the detected at-risk learners (i.e., FP and TP groups). In the latter groups, the time is computed before and after the learner is identified as a potential dropout. The largest and smallest access time values are provided by the TN group (i.e., correctly identified as non-at-risk learners) and TP group (i.e., correctly identified as at-risk learners), respectively. The number of learners wrongly detected as non-at-risk (i.e., FN) is significantly low in all CAA except the last one. We can observe that they are active in the classroom, similarly to learners in the TN group. Therefore, the DTWin approach is unable to detect them. However, the FP group is the relevant one that negatively impacts the TPR metric's performance because learners are detected as potential dropouts, but they finally submit the CAA. We can observe a relevant insight when comparing the access times before and after the at-risk alarm is triggered. Learners in the TP group have similar access times before and after the identification since the intervention mechanism does not impact their behavior. Conversely, FP learners have been impacted by the intervention mechanism, and their average time increases significantly near the values of the active learners, except in the CAA5.

Finally, Table 4 analyses the performance of the risk-level identification mechanism based on the decision tree provided in Fig. 2. The table summarizes the total number of learners (*n.*) identified for each risk level, the number of learners who submit the activity (*SUB*), and the percentage who submit compared to the total learners of the risk level (*%SUB*). Also, the percentages are highlighted in italic when the risk level has correctly identified more than 50% of the learners assigned to such risk level. Otherwise, the percentages are highlighted in bold. As we can observe, the problem with the CAA5 due to the different types of learners is propagated to the identification mechanism. When observing the risk levels, the G risk level correctly identifies non-at-risk students. The Y30% subsumes those learners that they are at 30% of the DTWin size. In this case, the performance is lower since some learners who finally have not submitted the CAA decrease their engagement within the online classroom near the end of the CAA duration. However, the system does not have time to raise the dropout alarm (i.e., there are not enough days left to reach the DTWin size). We can also observe that some learners who have not submitted some previous CAA are still active (YSNS risk level). The performance is lower than Y30% risk level since not submitting a CAA is a demotivating event that impacts engagement to continue the course.

Concerning NS and NS2 risk levels mostly identify learners who finally decided to drop out of the course. Also, the mechanism does not identify any active learner after not submitting more than one CAA (i.e., YSNS2). Thus, NS2 learners are the materialized dropouts, which justifies that an intervention message is unnecessary for this risk level. YLowD and R levels are activated depending on whether the DTWin's TPR metric is lower or higher than the 70% threshold to consider a high-quality model. In these cases, wrong identifications are mostly in CAA1 and CAA2. Here, we can also see the insight observed in Table 3 by risk level. Learners that increase engagement after at-risk identification are clustered in YLowD and R levels. That is, learners with low engagement at the beginning of the CAA become more active after receiving the intervention message.

Table 4 Performance of the risk-level identification on semester pilot

Activity		G	Y30%	YSNS	YSNS2	YLowS	YLowD	R	NS	NS2
CAA1	n	135	258	0	0	0	188	0	0	0
	SUB	131	247	0	0	0	172	0	0	0
	%SUB	<i>97.04</i>	<i>95.74</i>	0	0	0	91.49	0	0	0
CAA2	n	361	92	1	0	0	0	94	33	0
	SUB	315	65	1	0	0	0	70	11	0
	%SUB	<i>87.26</i>	<i>70.65</i>	<i>100</i>	0	0	0	74.47	<i>33.33</i>	0
CAA3	n	478	2	13	0	0	0	4	61	23
	SUB	449	2	8	0	0	0	1	23	0
	%SUB	<i>93.93</i>	<i>100</i>	<i>61.54</i>	0	0	0	<i>25.00</i>	<i>37.70</i>	0
CAA4	n	475	3	14	0	0	0	1	19	69
	SUB	437	2	7	0	0	0	1	6	1
	%SUB	<i>92.00</i>	<i>66.67</i>	<i>50.00</i>	0	0	0	100	<i>31.58</i>	<i>1.45</i>
CAA5	n	296	89	9	0	0	20	0	73	94
	SUB	161	8	8	0	0	3	0	45	0
	%SUB	<i>54.39</i>	8.99	<i>88.89</i>	0	0	<i>15.00</i>	0	61.64	0

Table 5 Boschloo's unconditional test for each activity

	CAA1	CAA2	CAA3	CAA4	CAA5
P-value	2.20e ⁻¹⁶	3.31e ⁻⁰⁹	6.29e ⁻⁰⁹	2.77e ⁻⁰⁶	0.036

Table 6 Dropout by each group of participants

	Signed (%)	Not signed (%)	Previous semester (%)
CAA1	5.34	24.20	16.94
CAA2	20.48	37.77	25.27
CAA3	16.87	32.98	22.43
CAA4	21.86	35.37	27.42
CAA5	61.27	67.02	63.76
End course	18.68	30.00	23.65
n	581	376	1036

RQ2. Is dropout decreased when using the LIS system?

We are interested in determining whether the intervention mechanism impacts the course dropout. In the previous section, we observed an engagement increment, but such increment is not enough evidence of dropout reduction. First, a statistical hypothesis testing with Boschloo's unconditional test is performed to see the association between submitting the CAA and participating in the pilot. Concretely, no association would imply that participating in the pilot and receiving information from the LIS system has no impact on the submitting event. The two analyzed variables are participating in the pilot (i.e., signing the consent form) and submitting the CAA. The null hypothesis is that not submitting the CAA is more significant in the learners who participated in the pilot. P-value results of the Boschloo's test are summarized in Table 5. The null hypothesis can be rejected in all the CAA, even CAA5.

After the statistical testing, the dropout is analyzed for each CAA and at the end of the course in Table 6. The dropout is shown for the different groups: participating in the pilot (signed), not participating (not signed), and results of the previous semester (previous semester). The dropout rate for learners who signed is significantly lower than for the not signed group and the previous semester. Recall that learners from the previous semester are the ones who signed the consent and already obtained information about course failure. This may suggest that the intervention mechanism has impacted potential dropout learners' decision to submit the CAA.

Note that we obviated to analyze performance on the final mark distribution because the result can be easily deduced from Table 6. Fewer dropout learners imply less low extreme values on the mark distribution (i.e., learners with a 0 final mark). Therefore, the median of the mark distribution will be better in the signed group.

RQ3. Is there any relationship between dropout and failure risk-level identification mechanisms?

Finally, we explore the relationship between both risk-level identification mechanisms. Since the LIS system has two predictive models and provides an intervention mechanism

for each type of risk, we are interested in analyzing the similarity between both mechanisms. Assigning similar risk levels would imply that the mechanisms are overlapping; therefore, the models predict the risk of course failure and dropout as the same outcome.

Here, we define similarity as learners obtaining the same risk levels in both models. Two statistical tests have analyzed the similarity, and the results are summarized in Table 7. First, a Chi-squared test is performed to check the independence between the two variables (i.e., risk-level identification mechanisms). In the case of dependency, the Chi-squared test does not give enough evidence about the meaningful correlation among variables. Therefore, the Cramer’s V test is done to measure the percentual association between both variables.

For the Chi-squared test, the null hypothesis is that there is no difference between both mechanisms. The hypothesis cannot be rejected in any CAA. Thus, there is some association that is thoroughly analyzed by the Cramer’s V test. The association varies depending on the CAA. Complete results can be reviewed in Appendix A, where each CAA’s contingency tables are shown. There is a low association in the CAA1 because there are fewer risk levels (i.e., NS and NS2 dropout risk cannot be still identified). However, the other CAA have a moderate association between 0.42 and 0.50. The similarity between mechanisms is produced in the similar low-risk levels (i.e., G, GLow, and Y30%) and the NS2 risk level that implies the dropout has been materialized. Learners identified on such levels produce similar results in both models. Low-risk learners are active learners who submit CAA and will likely pass the course. Learners in the NS2 level are the learners that drop out from the course, and the correlation increases through the course. The differences are actually in at-risk levels of each intervention mechanism where the corresponding message can impact. Many of the learners identified as at-risk of dropping out (i.e., the R and YLowD levels) are finally submitting the CAA (i.e., they have a risk level different than NS in the risk level of failure). Learners who have not submitted a previous CAA and have been identified at risk of dropping out are also distributed among different risk levels of failure. This may be by the nature of the flexible assessment model (i.e., it is still possible to be eligible for the VT by submitting the remaining CAA) or by the guidance of the intervention messages. Also, we observe that learners identified as medium and at-risk levels of failure (i.e., YPassActivity, YNotPassLow, and Red) come from non-at-risk dropout levels. Thus, learners who submit the CAA but have difficulties obtaining the minimum grade predicted by the course failure model can be helped by the failure intervention mechanism. However, we can observe in Appendix A that they are still some learners not identified as potential dropout learners (i.e., G level) that finally have not submitted the CAA (i.e., NS risk level of failure). The dropout risk-level identification mechanism cannot detect them during the CAA (i.e., they are the non-at-risk incorrectly identified group in Table 2).

Table 7 Correlation analysis between course dropout and failure risk-level identification mechanisms

	CAA1	CAA2	CAA3	CAA4	CAA5
P-value	0.99	0.99	1.00	1.00	0.99
Chi-squared					
Cramer’s V	0.25	0.48	0.42	0.42	0.50
Association	Low	Moderate	Moderate	Moderate	Moderate

Cramer’s V levels: 0–0.25: Low, 0.26–0.50: Moderated, 0.51–1.00: High

Conclusions, limitations, and future research

The insights collected in the Results section allow us to answer the research questions. Concerning RQ1. How accurate is the dropout risk-level identification on the LIS system?, we have shown the accuracy of the DTWin approach on the validation test (Table 1) and within the semester pilot (Table 2). The DTWin approach successfully detects most non-at-risk learners. The combination of data from the learner's profile, performance, and clickstream contributes to knowing the learners' situation at any moment. Profile information and performance allow assessing the learner's cognitive level, and the clickstream helps to evaluate the learner's engagement. Engagement can be considered a manifestation of motivation (Hew, 2015). The approach has some difficulties in the CAA1 because performance data (i.e., grades from previous CAA) are unavailable. Such data are relevant to identifying learners' status (Mubarak et al., 2020). The DTWin size for the CAA5 also has a low accuracy due to the characteristics of the assessment model applied in the piloted course.

The performance detecting non-at-risk dropout learners is similar to the validation test. However, we observe a significantly lower accuracy in detecting dropout at-risk learners, which is also manifested in the performance of the at-risk identification mechanism (Table 4). This lower identification is caused by the efficacy of the combination of the identification and intervention mechanism to engage learners. As observed in Table 3, access time in the online classroom substantially increases for learners stated as "at-risk incorrectly identified (FP)." This effect can be merely by triggering the at-risk alarm. Learners are informed within the dashboards about the potential at-risk situation. However, learners who do not access the VLE are notified by the intervention mechanism that activates them. Messages have been designed to increase motivation by setting a short-term goal. Learners are encouraged to submit the ongoing CAA by informing them of the potential negative consequences of non-submitting it. Giving learners recommendations about time management and a specific goal rather than telling them to do their best increases motivation (Locke & Latham, 2002), satisfaction (Henry, 2018), self-regulation (Veletsianos et al., 2021), and efficacy (Latham & Locke, 2007). Additionally, teachers can provide supplementary learning materials and exercises when necessary to help learners reach their goals. Such additional resources or exercises to acquire knowledge help learners self-regulate or set shorter self-goals. Goal-directed behavior promotes responsibility (Elliot & Fryer, 2008) and, in the end, also affects motivation (Zimmerman, 1990). Additionally, the dashboard design with predictive and descriptive information is underpinned by the Self-Regulated Learning theory that might support learners' independent learning, self-efficacy, self-regulation, and awareness of their progress (Jivet et al., 2020).

Concerning RQ2. Is dropout decreased when using the LIS system?, the LIS system successfully impacts learners' performance in reducing the dropout issues. The dropout in participating learners has decreased significantly in all CAA, with a relevant 12% difference between participants and not participants at the end of the course and a 5% difference compared with the previous semester with only the course failure mechanism. Thus, interventions increase retention throughout the course, as other authors pointed out (Borrella et al., 2019; Boudjehem & Lafifi, 2021; NeCamp et al., 2019; Xavier & Meneses, 2022). Learners not participating in the pilot receive only feedback after CAA

within the VLE when they fail or upon request. Learners may feel unmotivated considering a support deficit from the teacher (Hone & el Said, 2016). This may cause insecurity about passing the course or feeling overwhelmed about the following activities or competencies to acquire. The literature reports that isolation and poor feedback are reasons for not continuing to engage across the online course (Bakar et al., 2020; Ross & McNealy, 2020).

Finally, RQ3. Is there any relationship between dropout and failure risk-level identification mechanisms?, we claim that it is crucial to discern between both risk-level identification mechanisms. Some works are not distinguishing between them (Boudjehem & Lafifi, 2021; Guerrero-Roldán et al., 2021), and the interventions may not be correctly aligned with the actual risk. Such approaches consider that the risk of course failure subsumes dropout risk issues and constraints the early identification of dropout learners since the dropout may be already materialized when detected (Yair et al., 2020). Table 7 and Appendix A show a moderate correlation between both identification mechanisms. The results are expected since low-risk (i.e., G, GLow, Y30% levels) and materialized NS2 levels are similar for active learners and learners with difficulties. We also observed that the differences are actually in at-risk levels of each intervention mechanism where the corresponding message can impact. Thus, the timely intervention of the dropout mechanism is essential to impact the potential dropout learners successfully. In contrast, the course failure mechanism is crucial to help active learners who have difficulties passing the course. The DTWin approach provides a dynamic interval period of variable size that allows an intervention for each CAA which improves the number of interventions per course regarding previous approaches. The dynamic interval approach is significantly different from other approaches with a limited intervention during the course (Borrella et al., 2019; Burgos et al., 2018; Figueroa-Cañas & Sancho-Vinuesa, 2021) or weekly intervals (NeCamp et al., 2019) that may cause misprediction learners' risk status. Note that sending messages to the wrong audience may create untrustworthiness about the system and demoralize learners (Woodley & Simpson, 2013).

This study has some limitations. First, there is a self-selection bias induced by the research design imposed by the Ethical Committee. The learners who consented to participate tend to be active, engaged, and motivated to test innovative learning tools. Second mortality bias on non-participating learners may significantly affect such group's dropout performance. However, we can observe that despite such limitations, the interventions have positively impacted detected dropout learners mitigating the dropout at-risk situations.

In future work, we will dig into learners' and teachers' opinions about the utilization of the system. It is relevant to know the motivation and beliefs for using the EWS and its perceived usefulness for achieving their learning goals. Also, we are interested in understanding the difference between the different risk levels and how this differentiation may enlighten how messages must be constructed to address learners' difficulties better. Finally, the system personalization highly depends on trained models and teachers' messages. Models predict failure and the likelihood of dropping out, but there is no information about the no acquisition of which skills, competences, or knowledge may produce such events. Therefore, we propose to explore more specific models trained by skills or competence assessment to know which ones are associated with each risk

event. Additionally, such models can be used to create more specific interventions to help learners achieve critical skills or competence to pass a course or avoid dropping out (Bartimote-Aufflick et al., 2016). Such enhancements can help to build a more personalized system tailored to individual problems that learners may have.

Appendix A

Table 8 summarizes the contingency table between the course dropout and failure risk-level identification mechanisms. The row and columns represent the risk levels for course failure and for dropping out, respectively. The interrelation between risk

Table 8 Contingency table between course dropout and failure risk-level identification mechanism

Risk of failure	Risk of dropping out					
	G	Y30%	YSNS	YLowD	NS	NS2
CAA1						
G	26.60%	47.78%	0.00%	25.62%	0.00%	0.00%
GLow	14.89%	37.59%	0.00%	47.52%	0.00%	0.00%
YPassActivity	21.05%	31.58%	0.00%	47.37%	0.00%	0.00%
YNotPassLow	0.00%	60.00%	0.00%	40.00%	0.00%	0.00%
NS	20.00%	20.00%	0.00%	60.00%	0.00%	0.00%
CAA2						
G	76.82%	14.09%	0.00%	9.09%	0.00%	0.00%
GLow	69.02%	14.67%	0.00%	16.30%	0.00%	0.00%
YPassActivity	33.33%	0.00%	11.11%	11.11%	44.44%	0.00%
Red	38.18%	14.55%	0.00%	36.36%	10.91%	0.00%
NS	45.56%	28.89%	0.00%	25.56%	0.00%	0.00%
NS2	0.00%	0.00%	0.00%	0.00%	100.00%	0.00%
CAA3						
G	100.00%	0.00%	0.00%	0.00%	0.00%	0.00%
GLow	97.14%	2.86%	0.00%	0.00%	0.00%	0.00%
YPassActivity	56.58%	1.32%	10.53%	1.32%	30.26%	0.00%
Red	100.00%	0.00%	0.00%	0.00%	0.00%	0.00%
NS	92.59%	0.00%	0.00%	7.41%	0.00%	0.00%
NS2	5.71%	0.00%	7.14%	0.00%	54.29%	32.86%
CAA4						
G	99.65%	0.35%	0.00%	0.00%	0.00%	0.00%
GLow	100.00%	0.00%	0.00%	0.00%	0.00%	0.00%
YPassActivity	76.19%	0.00%	14.29%	0.00%	9.52%	0.00%
Red	75.00%	4.17%	4.17%	4.17%	8.33%	4.17%
NS	95.45%	4.55%	0.00%	0.00%	0.00%	0.00%
NS2	5.38%	0.00%	7.53%	0.00%	13.98%	73.12%
CAA5						
G	73.55%	4.52%	3.23%	0.65%	18.06%	0.00%
GLow	61.42%	28.74%	0.39%	4.72%	4.72%	0.00%
YPassActivity	54.55%	0.00%	18.18%	9.09%	18.18%	0.00%
Red	64.29%	0.00%	0.00%	0.00%	35.71%	0.00%
NS	40.00%	36.00%	0.00%	24.00%	0.00%	0.00%
NS2	0.82%	0.00%	0.82%	0.00%	21.31%	77.05%

levels is shown in distribution percentages for each risk of failure, i.e., the sum of the values for each row is 100%. The codes for the risk for failure are G for low risk, GLow for low risk and not passed CAA, YPassActivity for medium risk with the grade smaller than the prediction but activity passed, YNotPassLow for high risk but the accuracy of the predictive model (i.e., TPR) below 70%, Red for high risk and accuracy of the predictive model above 70%, and NS and NS2 not submitted one and more than one activity, respectively.

Abbreviations

LIS	Learning intelligent system
VLE	Virtual learning environment
HE	Higher education
AI	Artificial intelligence
MOOC	Massive open online course
EWS	Early warning system
CA	Continuous assessment
IT	Information technology
CAA	Continuous assessable activity
VT	Validation test
FM	Final mark
GDPR	General data protection regulation
PDAR	Profiled dropout at risk
DTWin	Dropout temporal window

Acknowledgements

We thank the eLearning Innovation Center (eLinC) and the eLearning Data Analysts unit at the Universitat Oberta de Catalunya for the support in the access and utilization of the institutional data mart.

Author contributions

All authors conceptualized the research design and methodology. AEGR and MERG contributed to related work revision, interpretation, and writing. DB contributed to software development, data collection, analysis and interpretation, and writing. PC contributed to the piloting process, analysis, and writing. All authors read and approved the final manuscript.

Funding

This research has been funded by the eLearn Center at Universitat Oberta de Catalunya through the project New Goals 2018NG001 "LIS: Learning Intelligent System."

Availability of data and materials

Data from the LIS system include personal identifying information when merged with information from the signed consent form, so they cannot be disclosed to the general public.

Declarations

Ethics approval and consent to participate

The UOC Research Ethical Committee requires learners to give their explicit consent to be included in any study. A consent form was required for all participants prior to starting, and Ethical Committee approved the research procedure.

Consent for publication

All authors consent to its publication. A consent form was signed by all participants involved in this research paper.

competing interests

The authors declare that they have no competing interests.

Received: 27 July 2022 Accepted: 26 October 2022

Published online: 10 January 2023

References

- Bağrıncak Yılmaz, A., & Karataş, S. (2022). Why do open and distance education students drop out? Views from various stakeholders. *International Journal of Educational Technology in Higher Education*, 19(1), 1–22. <https://doi.org/10.1186/S41239-022-00333-X>
- Bakar, A., Shah, K., & Qingyu, X. (2020). The effect of communication barriers on distance learners achievements. *Revista Argentina De Clínica Psicológica*, 29(5), 248–264. <https://doi.org/10.24205/03276716.2020.1027>
- Bandura, A. (1997). Self-efficacy: The exercise of control. *Choice Reviews Online*. <https://doi.org/10.5860/choice.35-1826>
- 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 (switzerland)*, 10(13), 4427. <https://doi.org/10.3390/app10134427>

- Bartimote-Aufflick, K., Bridgeman, A., Walker, R., Sharma, M., & Smith, L. (2016). The study, evaluation, and improvement of university student self-efficacy. *Studies in Higher Education*. <https://doi.org/10.1080/03075079.2014.999319>
- Borrella, I., Caballero-Caballero, S., & Ponce-Cueto, E. (2019). Predict and intervene: Addressing the dropout problem in a MOOC-based program. *Proceedings of the 6th 2019 ACM Conference on Learning at Scale, L@S 2019*. <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>
- Boudjehem, R., & Lafifi, Y. (2021). A new approach to identify dropout learners based on their performance-based behavior. *Journal of Universal Computer Science*. <https://doi.org/10.3897/jucs.74280>
- Broadbent, J., & Poon, W. L. (2015). Self-regulated learning strategies & academic achievement in online higher education learning environments: A systematic review. *In Internet and Higher Education*. <https://doi.org/10.1016/j.iheduc.2015.04.007>
- Burgos, C., Campanario, M. L., de la Peña, D., Lara, J. A., Lizcano, D., & Martínez, M. A. (2018). Data mining for modeling students' performance: A tutoring action plan to prevent academic dropout. *Computers & Electrical Engineering*, 66, 541–556. <https://doi.org/10.1016/J.COMPELECENG.2017.03.005>
- Chen, Y., Chen, Q., Zhao, M., Boyer, S., Veeramachaneni, K., & Qu, H. (2017). DropoutSeer: Visualizing learning patterns in Massive Open Online Courses for dropout reasoning and prediction. *2016 IEEE Conference on Visual Analytics Science and Technology, VAST 2016 - Proceedings*. <https://doi.org/10.1109/VAST.2016.7883517>
- Dalipi, F., Imran, A. S., & Kastrati, Z. (2018). MOOC dropout prediction using machine learning techniques: Review and research challenges. *IEEE Global Engineering Education Conference, EDUCON, 2018-April*, 1007–1014. <https://doi.org/10.1109/EDUCON.2018.8363340>
- Darkenwald, G. G., & Gavin, W. J. (1987). Dropout as a function of discrepancies between expectations and actual experiences of the classroom social environment. *Adult Education Quarterly*. <https://doi.org/10.1177/0001848187037003003>
- de Barba, P. G., Malekian, D., Oliveira, E. A., Bailey, J., Ryan, T., & Kennedy, G. (2020). The importance and meaning of session behaviour in a MOOC. *Computers and Education*. <https://doi.org/10.1016/j.compedu.2019.103772>
- Douglas, K. A., Merzdorf, H. E., Hicks, N. M., Sarfraz, M. I., & Bermel, P. (2020). Challenges to assessing motivation in MOOC learners: An application of an argument-based approach. *Computers and Education*. <https://doi.org/10.1016/j.compedu.2020.103829>
- Dourado, R. A., Rodrigues, R. L., Ferreira, N., Mello, R. F., Gomes, A. S., & Verbert, K. (2021). A teacher-facing learning analytics dashboard for process-oriented feedback in online learning. *ACM International Conference Proceeding Series*. <https://doi.org/10.1145/3448139.3448187>
- Elliot, A. J., & Fryer, J. W. (2008). The goal construct in psychology. *Handbook of Motivation Science*, 18, 235–250.
- El-Sabagh, H. A. (2021). Adaptive e-learning environment based on learning styles and its impact on development students' engagement. *International Journal of Educational Technology in Higher Education*. <https://doi.org/10.1186/s41239-021-00289-4>
- Eriksson, T., Adawi, T., & Stöhr, C. (2017). "Time is the bottleneck": a qualitative study exploring why learners drop out of MOOCs. *Journal of Computing in Higher Education*. <https://doi.org/10.1007/s12528-016-9127-8>
- Figueroa-Cañas, J., & Sancho-Vinuesa, T. (2021). Changing the recent past to reduce ongoing dropout: An early learning analytics intervention for an online statistics course. *Open Learning*. <https://doi.org/10.1080/02680513.2021.1971963>
- Goel, Y., & Goyal, R. (2020). On the effectiveness of self-training in MOOC dropout prediction. *Open Computer Science*, 10(1), 246–258. <https://doi.org/10.1515/comp-2020-0153>
- Grau-Valldosera, J., & Minguillón, J. (2014). Rethinking dropout in online higher education: The case of the universitat oberta de catalunya. *International Review of Research in Open and Distance Learning*, 15(1), 290–308. <https://doi.org/10.19173/irrodl.v15i1.1628>
- Greenland, S. J., & Moore, C. (2022). Large qualitative sample and thematic analysis to redefine student dropout and retention strategy in open online education. *British Journal of Educational Technology*. <https://doi.org/10.1111/bjet.13173>
- Guerrero-Roldán, A. E., Rodríguez-González, M. E., Bañeres, D., Elasi-Ejjaberi, A., & Cortadas, P. (2021). Experiences in the use of an adaptive intelligent system to enhance online learners' performance: a case study in Economics and Business courses. *International Journal of Educational Technology in Higher Education*. <https://doi.org/10.1186/s41239-021-00271-0>
- Hart, C. (2012). Factors associated with student persistence in an online program of study: A review of the literature. *Journal of Interactive Online Learning* 11(1).
- Henry, M. (2018). The online student experience: An exploration of first-year university students' expectations, experiences and outcomes of online education. *In The Grants Register 2020*.
- Hew, K. F. (2015). Towards a model of engaging online students: Lessons from MOOCs and four policy documents. *International Journal of Information and Education Technology*. <https://doi.org/10.7763/ijiet.2015.v5.543>
- Hone, K. S., & el Said, G. R. (2016). Exploring the factors affecting MOOC retention: A survey study. *Computers and Education*. <https://doi.org/10.1016/j.compedu.2016.03.016>
- Itani, A., Brisson, L., & Garlatti, S. (2018). Understanding Learner's Drop-Out in MOOCs. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 11314 LNCS. https://doi.org/10.1007/978-3-030-03493-1_25
- Jivet, I., Scheffel, M., Schmitz, M., Robbers, S., Specht, M., & Drachsler, H. (2020). From students with love: An empirical study on learner goals, self-regulated learning and sense-making of learning analytics in higher education. *Internet and Higher Education*. <https://doi.org/10.1016/j.iheduc.2020.100758>
- Kizilcec, R. F., Pérez-Sanagustín, M., & Maldonado, J. J. (2016). Recommending self-regulated learning strategies does not improve performance in a MOOC. *L@S 2016—Proceedings of the 3rd 2016 ACM Conference on Learning at Scale*. <https://doi.org/10.1145/2876034.2893378>
- Kuechler, W., & Vaishnavi, V. (2012). A framework for theory development in design science research: Multiple perspectives. *Journal of the Association for Information Systems*. <https://doi.org/10.17705/1jais.00300>

- Kurtz, G., Kopolovich, O., Segev, E., Sahar-Inbar, L., Gal, L., & Hammer, R. (2022). Impact of an instructor's personalized email intervention on completion rates in a massive open online course (MOOC). *Electronic Journal of E-Learning*, 20(3), 325–335.
- Latham, G. P., & Locke, E. A. (2007). New developments in and directions for goal-setting research. *European Psychologist*. <https://doi.org/10.1027/1016-9040.12.4.290>
- Lee, Y., & Choi, J. (2011). A review of online course dropout research: Implications for practice and future research. In *Educational Technology Research and Development*. <https://doi.org/10.1007/s11423-010-9177-y>
- Locke, E. A., & Latham, G. P. (2002). Building a practically useful theory of goal setting and task motivation: A 35-year odyssey. *American Psychologist*. <https://doi.org/10.1037/0003-066X.57.9.705>
- Minguillón, J., Conesa, J., Rodríguez, M. E., & Santanach, F. (2018). Learning analytics in practice: Providing E-learning researchers and practitioners with activity data. In J. M. Spector, V. Kumar, A. Essa, Y. M. Huang, R. Koper, R. A. W. Tortorella, T. W. Chang, Y. Li, & Z. Zhang (Eds.), *Frontiers of cyberlearning: emerging technologies for teaching and learning* (pp. 145–167). Springer. https://doi.org/10.1007/978-981-13-0650-1_8
- Moreno-Marcos, P. M., Alario-Hoyos, C., Munoz-Merino, P. J., & Kloos, C. D. (2019). Prediction in MOOCs: A review and future research directions. *IEEE Transactions on Learning Technologies*. <https://doi.org/10.1109/TLT.2018.2856808>
- Mubarak, A. A., Cao, H., & Zhang, W. (2020). Prediction of students' early dropout based on their interaction logs in online learning environment. *Interactive Learning Environments*. <https://doi.org/10.1080/10494820.2020.1727529>
- NeCamp, T., Gardner, J., & Brooks, C. (2019). Beyond A/B testing: Sequential randomization for developing interventions in scaled digital learning environments. *ACM International Conference Proceeding Series*. <https://doi.org/10.1145/3303772.3303812>
- Oates, B. J. (2006). Researching information systems and computing. In *Inorganic chemistry* (Vol. 37). SAGE Publications Ltd. <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2836698&tool=pmcentrez&rendertype=abstract>
- Park, J. H., & Choi, H. J. (2009). Factors influencing adult learners' decision to drop out or persist in online learning. *Educational Technology and Society*, 12(4).
- Rodríguez, M. E., Guerrero-Roldán, A. E., Baneres, D., & Karadeniz, A. (2019). Towards an intervention mechanism for supporting learners performance in online learning. *ICERI2019 Proceedings*, 5136–5145. <https://doi.org/10.21125/iceri.2019.1244>
- Rodríguez, M. E., Guerrero-Roldán, A. E., Baneres, D., & Karadeniz, A. (2022). An Intelligent Nudging System to Guide Online Learners. *International Review of Research in Open and Distance Learning*. <https://doi.org/10.19173/IRRODL.V22I4.5407>
- Ross, E., & McNealy, K. (2020). *Creating connection between virtual learners*. https://econometricainc.com/wp-content/uploads/2020/10/Polling_Blue-Paper_Oct2020.pdf
- Sedrakyan, G., Malmberg, J., Verbert, K., Järvelä, S., & Kirschner, P. A. (2020). Linking learning behavior analytics and learning science concepts: Designing a learning analytics dashboard for feedback to support learning regulation. *Computers in Human Behavior*. <https://doi.org/10.1016/j.chb.2018.05.004>
- Simpson, C., Baker, K., & Mellinger, G. (1980). Conventional failures and unconventional dropouts: comparing different types of university withdrawals. *Sociology of Education*. <https://doi.org/10.2307/2112529>
- Stephen, J. S., Rockinson-Szapkiw, A. J., & Dubay, C. (2020). Persistence model of non-traditional online learners: Self-efficacy, self-regulation, and self-direction. *American Journal of Distance Education*. <https://doi.org/10.1080/08923647.2020.1745619>
- Stone, C., & O'Shea, S. (2019). Older, online and first: Recommendations for retention and success. *Australasian Journal of Educational Technology*, 35(1), 57–69. <https://doi.org/10.14742/ajet.3913>
- Tang, J. K. T., Xie, H., & Wong, T. L. (2015). A big data framework for early identification of dropout students in MOOC. *Communications in Computer and Information Science*. https://doi.org/10.1007/978-3-662-48978-9_12
- Teusner, R., Hille, T., & Staubitz, T. (2018). Effects of automated interventions in programming assignments: Evidence from a field experiment. *Proceedings of the 5th Annual ACM Conference on Learning at Scale, L at S 2018*. <https://doi.org/10.1145/3231644.3231650>
- Thaler, R. H., & Sunstein, C. R. (2008). Nudge: Improving decisions about health, wealth, and happiness. *Nudge: Improving Decisions about Health, Wealth, and Happiness*. [https://doi.org/10.1016/s1477-3880\(15\)30073-6](https://doi.org/10.1016/s1477-3880(15)30073-6)
- Thalhammer, V., Hoffmann, S., von Hippel, A., & Schmidt-Hertha, B. (2022). Dropout in adult education as a phenomenon of fit—An integrative model proposal for the genesis of dropout in adult education based on dropout experiences. *European Journal for Research on the Education and Learning of Adults*.
- Tinto, V. (1975). Dropout from higher education: A theoretical synthesis of recent research. *Review of Educational Research*, 45(1), 89–125. <https://doi.org/10.3102/00346543045001089>
- Veletsianos, G., Kimmons, R., Larsen, R., & Rogers, J. (2021). Temporal flexibility, gender, and online learning completion. *Distance Education*. <https://doi.org/10.1080/01587919.2020.1869523>
- Whitehill, J., Mohan, K., Seaton, D., Rosen, Y., & Tingley, D. (2017). MOOC dropout prediction: How to measure accuracy? *L@S 2017—Proceedings of the 4th (2017) ACM Conference on Learning at Scale*, 161–164. <https://doi.org/10.1145/3051457.3053974>
- Woodley, A., & Simpson, O. (2013). Student Dropout: the elephant in the room. In O. Zawacki-Richter & T. Anderson (Eds.), *Online distance education: towards a research agenda* (pp. 459–483). AU Press, Athabasca University. <https://doi.org/10.15215/aupress/9781927356623.01>
- Xavier, M., & Meneses, J. (2020). A Literature Review on the Definitions of Dropout in Online Higher Education. *EDEN Conference Proceedings*, 0(1), 73–80. <https://doi.org/10.38069/edenconf-2020-ac0004>
- Xavier, M., & Meneses, J. (2022). Persistence and time challenges in an open online university: A case study of the experiences of first-year learners. *International Journal of Educational Technology in Higher Education*, 19(1), 31. <https://doi.org/10.1186/s41239-022-00338-6>
- Xenos, M., Pierrakeas, C., & Pintelas, P. (2002). A survey on student dropout rates and dropout causes concerning the students in the Course of Informatics of the Hellenic Open University. *Computers and Education*. [https://doi.org/10.1016/S0360-1315\(02\)00072-6](https://doi.org/10.1016/S0360-1315(02)00072-6)

- Yair, G., Rotem, N., & Shustak, E. (2020). The riddle of the existential dropout: Lessons from an institutional study of student attrition. *European Journal of Higher Education*, 10(4), 436–453. <https://doi.org/10.1080/21568235.2020.1718518>
- Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education—Where are the educators? *International Journal of Educational Technology in Higher Education*. <https://doi.org/10.1186/s41239-019-0171-0>
- Zimmerman, B. J. (1990). Self-regulated learning and academic achievement: An overview. *Educational Psychologist*. https://doi.org/10.1207/s15326985ep2501_2

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- ▶ Convenient online submission
- ▶ Rigorous peer review
- ▶ Open access: articles freely available online
- ▶ High visibility within the field
- ▶ Retaining the copyright to your article

Submit your next manuscript at ▶ [springeropen.com](https://www.springeropen.com)
