education sciences

MDPI

Article

Securing Equitable Success in the First Year: A Whole-Institution Case Study of Data-Driven Early Warning and Critical AI Literacy

António Luís Lopes [1,*], David Miguel Rodrigues [2], Rosário Mauritti [2], Mariana Roque Ferreira [2] and Sónia Pintassilgo [2]

[1] Centro de Investigação em Ciências da Informação, Tecnologias e Arquitetura, Instituto Universitário de Lisboa (ISCTE-IUL), 1649-026 Lisboa, Portugal
[2] Centro de Investigação e Estudos de Sociologia, Instituto Universitário de Lisboa (ISCTE-IUL), 1649-026 Lisboa, Portugal; david_miguel_rodrigues@iscte-iul.pt (D.M.R.); rosario.mauritti@iscte-iul.pt (R.M.); roque_ferreira@iscte-iul.pt (M.R.F.); sonia.cardoso@iscte-iul.pt (S.P.)
* Correspondence: antonio.luis@iscte-iul.pt

Abstract

Widening access has transformed who enters higher education but not who thrives once inside: inequality resurfaces as first-year failure and dropout, while the rapid spread of generative artificial intelligence (AI) threatens to open new divides. This article presents a whole-institution case study of how one Portuguese public university, Iscte—University Institute of Lisbon, has pursued equitable success in the first year through a single strategy with two reinforcing arms: a reactive arm that detects at-risk students and responds with personalized support, and a proactive arm that builds a baseline of critical AI literacy that is mandatory on one campus and optional on the other. Drawing on institutional records from 2016/17 to 2024/25, it combines a rule-based early warning alarm system, a comparison of machine learning models, and descriptive evidence on the literacy initiatives. The findings show that academic risk is concentrated in identifiable groups, that flagged students who engaged with support re-enrolled at higher rates than those who did not (though not distinguishably so once students who had already formalized a withdrawal are set aside, a group that took up no support and left without exception), and that a common literacy baseline is operationally achievable. The findings also expose the limits of each, as support reaches only a minority of those flagged and exploratory predictive models transfer poorly to the following cohort. Together, these observational results frame reactive support and proactive capacity-building as facets of one inclusion strategy.

Keywords: higher education; first-year transition; student dropout; learning analytics; early warning systems; AI-based detection; AI literacy

Academic Editor: Thomas M. Smith

Received: 15 June 2026
Revised: 30 August 2026
Accepted: 4 September 2026
Published: 5 September 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

1. Introduction

Over the past three decades, higher education systems across Europe and beyond have expanded from elite to mass and, increasingly, are becoming more accessible to people with different backgrounds. This expansion has been driven by rising national demand and, more recently, by intensifying international mobility, with the combined effect of diversifying the student body along socio-economic, cultural, linguistic, and educational lines (A. F. Costa & Lopes, 2011; Martins et al., 2017; Oliveira et al., 2021). The

diversification of who enters higher education is a substantial achievement of equity-oriented policy. Yet, a growing body of research makes clear that access is a necessary but insufficient condition for equity: once admitted, students from under-represented backgrounds frequently encounter unequal conditions for success, and inequality re-emerges in differential patterns of academic failure, disengagement, and dropout (A. F. Costa et al., 2014; Martins et al., 2023). Equity agendas have therefore shifted from a narrow concern with widening participation at the point of entry toward a broader concern with equitable success throughout the student trajectory.

Within that trajectory, the first year occupies a decisive position. It is the period in which the highest concentration of failure and dropout occurs, when the academic and social demands of becoming a student are most acute, and in which the gap between prior preparation and institutional expectation is most consequential (Casanova et al., 2021; Coulon, 2017; Gil et al., 2021). Especially for students who relocate from another region or country, who combine study with paid work, who are the first in their family to attend university, or who arrive with weaker digital or linguistic preparation, the first year can compound a sense of disorientation that, if unaddressed, accumulates silently into withdrawal (Casanova et al., 2021; White & Canning, 2023). The first year is thus the juncture at which an institution's commitment to equitable success is most severely tested.

This article frames the contemporary equity problem in two complementary ways. The first is the long-recognized observation that certain groups are systematically more exposed to academic risk and that institutions can (and should) identify and support them before difficulties become irreversible. The second is newer and less settled: the rapid diffusion of generative artificial intelligence (GenAI) threatens to create fresh inequalities because students arrive with markedly unequal prior exposure to, confidence with, and critical understanding of these tools (Bozkurt, 2024; Tlili et al., 2023). Left to chance, GenAI may widen, rather than narrow, the gaps that equity policy seeks to close. An institution serious about equitable success must therefore act on both fronts at once.

This study examines how a single institution, Iscte—University Institute of Lisbon, a Portuguese public university, has attempted to do so through the IN-Iscte—A Space to Grow project: an integrated, equity-oriented response funded under the Portuguese Recovery and Resilience Plan. The program is read here not as a collection of separate projects but as a single strategy operating through two reinforcing mechanisms. The first mechanism, detect-and-support, combines data-driven early warning systems—a transparent rule-based alarm and a set of machine learning models built on the institutional information system—with personalized, negotiated intervention by a multidisciplinary team. The second mechanism, emerging-gap-levelling, builds critical AI literacy across the whole institution through structured faculty development and a universal first-year student course so that no group is left behind on a fast-emerging competency. The central thesis is that equity in contemporary higher education requires both a reactive arm, which identifies and supports those most at risk, and a proactive arm, which ensures that emerging competencies are distributed inclusively, and that these two arms are coherent facets of a single inclusion strategy rather than unrelated initiatives.

The contribution of the article is threefold. First, it offers an integrated empirical account of an institutional equity strategy. Second, it advances a two-mechanism conceptual framework that connects learning analytics and digital literacy agendas, which are usually treated separately under a common equity rationale. Third, it derives transferable implications for other institutions seeking to move from access to equitable success. The analysis is organized around four research questions:

- RQ1 asks who is at risk and how vulnerability is distributed across student groups;
- RQ2 asks whether early detection coupled with personalized intervention shifts academic outcomes;

- RQ3 asks how exploratory predictive models trained on pre-intervention cohorts behave when transferred to a subsequent cohort and what the observed degradation can and cannot be attributed to;
- RQ4 asks whether a common literacy baseline can be delivered at institutional scale, which is a precondition for it to function as an inclusion measure.

The remainder of the article is structured as follows. Section 2 develops the conceptual framework, synthesizing the literature on widening participation, the first-year transition, vulnerable groups in Portuguese higher education, learning analytics, and the emerging AI digital divide. Section 3 describes the institutional setting, the data, and the three operational components of the strategy. Section 4 presents the results, organized around the four research questions. Section 5 discusses the integrated equity model, the failure–dropout distinction, and the role of AI literacy as inclusion. Section 6 sets out the limitations of the study, and Section 7 concludes with transferable implications and directions for future work.

2. Theoretical Framework

The expansion of higher education has been accompanied by an evolution in how equity itself is understood. Early equity agendas concentrated on widening participation: removing financial, geographic, and informational barriers so that students from working-class, rural, first-generation, and minority backgrounds could enter institutions historically reserved for a privileged minority (Martins et al., 2017; Oliveira et al., 2021). In the Portuguese context, successive policies have substantially democratized access, expanding the network of institutions, diversifying entry routes, and supporting students through social-action scholarships (Martins et al., 2017; Oliveira et al., 2021). These measures changed the composition of the student body in ways that earlier generations of policy intended.

However, research has repeatedly shown that equalizing entry does not equalize outcomes. Students from disadvantaged backgrounds tend to enter with weaker prior academic preparation, less familiarity with the implicit codes of academic life, more fragile financial situations, and a higher likelihood of combining study with employment or caring responsibilities (Almeida et al., 2012; A. F. Costa et al., 2014; A. F. Costa & Lopes, 2011). These conditions translate into measurable differences in achievement, progression, and completion. The equity problem, in other words, migrates from the gateway to the pathway and reappears as an equity–success gap that access policy alone cannot close (A. F. Costa et al., 2014; Martins et al., 2023). Recognizing this, contemporary research and policy have shifted attention toward equitable success: the proposition that institutions bear responsibility not only for admitting a diverse student body but for creating the conditions in which that diversity can succeed (Martins et al., 2023). This reframing motivates the present study, which treats the prevention of failure and dropout, and the inclusive distribution of emerging competencies, as integral to an equity agenda rather than as separate operational concerns (Capucha et al., 2021; Martins et al., 2023).

2.1. The First-Year Transition

The transition from secondary to higher education is not merely a change of institution but a change of role. The student must learn the tacit conventions, expectations, and intellectual dispositions of academic life. Coulon's notion of *le métier d'étudiant* (the craft, or trade, of being a student) captures this process of affiliation, in which newcomers must learn to decode institutional rules and routines that insiders take for granted (Coulon, 2017). Those who fail to make this affiliation in the first weeks and months are disproportionately likely to disengage. The literature on academic and social integration similarly

emphasizes that early integration, both into the intellectual life of the course and into the social fabric of the institution, is among the strongest predictors of persistence (Casanova et al., 2021).

A particularly important and well-documented phenomenon is the reluctance of vulnerable students to seek help. First-generation and minority students, precisely those who might benefit most from support services, are frequently the least likely to use them because of feelings of inadequacy, unfamiliarity, or not belonging inhibit help-seeking (White & Canning, 2023). The consequence is that academic difficulties accumulate unnoticed, often invisibly to the very services designed to address them, until they manifest as failure or withdrawal (Casanova et al., 2021; White & Canning, 2023). This asymmetry (those most in need are least likely to ask) provides a strong rationale for proactive institutional identification of risk rather than reliance on students to present themselves.

2.2. Under-Represented and Vulnerable Groups in Portuguese Higher Education

The diversification of the Portuguese student body has brought into universities several groups whose circumstances place them at elevated risk. International students, in particular students from the Portuguese-speaking African countries (PALOP), face a combination of linguistic and cultural contrast, divergent prior educational backgrounds, and socio-economic integration challenges (Matias et al., 2023; Mauritti et al., 2023). Many also confront administrative obstacles: visa processing delays frequently mean that PALOP students arrive in Portugal after the academic year has begun, enrolling late and missing the first weeks or months of teaching and assessment (Mauritti et al., 2023). Relocated students (those whose family residence is far from the institution) must manage the practical and financial burdens of living away from home in a context where affordable student accommodation is scarce, and the disruption of relocation has been associated with elevated anxiety, depression, and stress (Massano-Cardoso et al., 2024). Working students and mature students, often concentrated in evening or post-work degrees, must reconcile study with employment and, frequently, with family responsibilities. Scholarship holders, while benefiting from financial support, face the additional pressure of progression thresholds that condition the renewal of that support from the second year onwards; entry support is itself awarded on a means test, so for first-year students, scholarship status reflects household circumstances rather than prior academic performance. Students with special educational needs encounter barriers that vary widely with the nature of their needs and the responsiveness of provision.

These groups do not experience risk uniformly, and the relationship between status and outcome is neither simple nor deterministic. Some characteristics that might be assumed to be disadvantages prove (in certain institutional contexts) to be associated with resilience or success. This heterogeneity is itself an argument against generic intervention and is in favor of differentiated, context-sensitive support, a point to which the analysis returns.

2.3. Learning Analytics and Early Warning Systems

The proposition that institutions should proactively identify at-risk students has been operationalized, over the past two decades, through the field of learning analytics and, more specifically, through early warning systems. The traditional approach to identification was manual: tutors, course directors, and services would notice struggling students over the course of the year and respond on a case-by-case basis. While valuable, manual identification is limited in scalability, consistency, and speed, and it is precisely the silent, non-help-seeking cases that it tends to miss (White & Canning, 2023). Data-driven systems promise to address these limitations by drawing on the administrative and academic data that institutions already hold.

Two broad approaches can be distinguished. Rule-based systems apply transparent, human-defined thresholds (such as the proportion of credits completed) to flag students for attention. Their virtues are interpretability, ease of communication to non-technical staff, and direct alignment with institutional definitions of progress. Their limitation is that they encode existing assumptions rather than discovering patterns. Machine learning systems, by contrast, learn predictive relationships from historical data and can, in principle, detect risk earlier and more accurately, including before academic results are available (Behr et al., 2020; E. B. Costa et al., 2017; Realinho et al., 2022; Rodrigues et al., 2025). A consistent finding in this literature is that early academic performance is among the strongest predictors of subsequent success, which explains why models incorporating first-semester results thus substantially outperform those relying on enrolment data alone (Hellas et al., 2018; Rodrigues et al., 2025). Nevertheless, prediction from enrolment-only data is valuable precisely because it enables intervention before difficulties emerge (Rodrigues et al., 2025).

The literature also documents well-known limitations that temper enthusiasm. Predictive models are sensitive to model drift: when the environment changes, for example, when a new support program alters the very outcomes that the model was trained to predict, performance can degrade, and models may require monitoring and re-tuning (Rodrigues et al., 2025). Class imbalance (at-risk students are a minority) complicates training and evaluation. False positives carry real costs, both in the misallocation of scarce support resources and in the potential for stigmatization. There is, further, an intervention capacity problem that is frequently underappreciated: a model that flags many students is useless, or even harmful, if the institution lacks the capacity to act on the flags, so the choice of the model must balance predictive performance against the feasibility of response. Finally, the ethics of flagging (questions of consent, anonymization, data protection, and the tension between targeted intervention and privacy) are central rather than peripheral concerns (Pargman & McGrath, 2021). An equity-oriented early warning system must therefore be designed not only for accuracy but for transparency, proportionality, and humane response.

Three further methodological hazards are well documented and bear directly on the analysis reported below. First, target leakage, in which a feature encodes information about the outcome that would not have been available at the moment of prediction, is a leading cause of irreproducible performance claims across the applied machine learning literature (Kapoor & Narayanan, 2023); administrative status fields that are updated over the course of the year are a common vehicle for it. Second, under strong class imbalance, accuracy, recall and F1 computed with the majority class as the positive class can appear high while conveying nothing about detection of the minority at-risk class, so precision–recall summaries and explicit trivial baselines are required for the figures to be interpretable. Third, aggregate performance can conceal systematically worse performance for particular student groups, which in an equity-oriented deployment would divert scarce support away from those who need it most; subgroup or slicing evaluation is therefore a substantive validity check rather than an optional refinement (Baker & Hawn, 2022; Gardner et al., 2019).

A fourth consideration is of a different kind, since it concerns not a flaw to be avoided but a structural property of prediction in settings where the prediction is acted upon. Where a model's output triggers an intervention, the outcome that subsequently materializes reflects both the underlying risk and the institutional response to it, so that a successful intervention necessarily degrades the apparent accuracy of the model that prompted it. Clinical informatics describes this as the self-defeating prognostic model, in which deployment triggers the very interventions that falsify the model's predictions (Lenert et al., 2019), and the machine learning literature formalizes it as performative

prediction, in which the act of predicting alters the distribution being predicted (Perdomo et al., 2020). The consequences for how a deployed early warning system should be evaluated are taken up in Sections 4.3 and 7.

2.4. Generative Artificial Intelligence and the Emerging Digital Divide

A second, newer equity concern has emerged with striking speed. The public release of conversational GenAI tools, most prominently ChatGPT in late 2022, brought to a mass audience systems capable of producing fluent, human-like text. Their diffusion has been extraordinarily rapid, and their implications for higher education are profound (Tlili et al., 2023). On one hand, generative AI offers genuine pedagogical potential: as a personalized tutor, a brainstorming partner, a source of individualized feedback, and a means of exploring alternative perspectives. On the other, it raises serious concerns about over-reliance, the erosion of cognitive effort and critical thinking, and new forms of academic dishonesty for which reliable detection does not exist (Heung & Chiu, 2025; Tlili et al., 2023).

For an equity agenda, the central risk is that GenAI reproduces and amplifies existing inequalities. Students do not arrive with equal prior access to these tools, equal confidence in using them, or equal critical understanding of their potential and limitations. Those who already possess digital and cultural capital are better positioned to exploit AI productively, while those who do not may either avoid it, to their disadvantage, or use it uncritically, to the detriment of their learning (Bozkurt, 2024). Thus, the capacity to understand how these systems work and fail, to evaluate the reliability of their outputs, and to use them ethically and as a support for rather than a substitute for one's own reasoning—denoted as AI literacy—becomes a new dimension of educational equity (Bozkurt, 2024). The research further emphasizes that this literacy does not develop spontaneously: unsupervised student use tends toward shallow, over-reliant engagement, whereas structured, teacher-mediated use can preserve and even strengthen critical thinking (Heung & Chiu, 2025; Ramos et al., 2025). The educator is therefore an indispensable mediator, and the redesign of assessment toward authentic tasks that foreground critical analysis and explicit reasoning is preferable to a posture of surveillance and detection (Mah & Groß, 2024; Tlili et al., 2023). Faculties, however, require their own development to take up this role, since they too vary in their familiarity, self-efficacy, and professional preparation for AI (Mah & Groß, 2024).

This concern is increasingly codified in European competence policy, which supplies the reference frame within which an institutional AI literacy offer should be read. DigComp 2.2 extended the European Digital Competence Framework for Citizens with explicit examples of the knowledge, skills and attitudes required to engage with AI systems (Vuorikari et al., 2022), and its successor DigComp 3.0 integrates AI competence transversally across all twenty-one competences rather than isolating it as a specialism (Cosgrove & Cachia, 2025). DigComp 3.0 also makes explicit a set of prerequisites for attaining even a basic level of digital competence, among them a stable internet connection, an adequate device, and access to technical assistance and guidance. That prerequisite structure matters for the equity argument advanced here because it establishes that digital and AI competence is distributed by material circumstance and not by aptitude alone, which is the same pattern that widening-participation policy encounters at the level of academic outcomes. It follows that a policy of equal opportunity in digital and AI literacy cannot consist in equal availability of provision. Where provision is optional, availability reproduces the distribution of prior confidence and cultural capital. Only compulsory or actively compensatory provision addresses the catch-up problem that the frameworks identify. UNESCO's AI competency framework for students converges on a comparable position, organizing twelve competencies around a human-centered mindset, ethics,

techniques and system design across three progression levels (UNESCO, 2024). In the European regulatory context, Article 4 of the AI Act, applicable since February 2025, further obliges providers and deployers of AI systems to ensure a sufficient level of AI literacy among their staff and others operating such systems on their behalf (European Union, 2024), which gives universities a compliance rationale alongside the educational one. Read together, these instruments frame AI literacy as a general entitlement instead of a specialist skill, and they support reading a first-year literacy baseline as an equity measure rather than a technical addition to the curriculum.

2.5. Synthesis

The work reviewed above is usually pursued separately: learning analytics and dropout prevention in one tradition, digital and AI literacy in another. This article argues that they are best understood as two mechanisms of a single equity logic. The first mechanism is reactive: because vulnerable students are unevenly distributed and unlikely to seek help, the institution must detect risk proactively and respond with differentiated, personalized support. The second mechanism is proactive: because emerging competencies are unequally distributed at entry, the institution must build a common literacy baseline so that technological change narrows rather than widens gaps. Both mechanisms share the same equity rationale, and both are concentrated where they matter most, in the first-year transition. The gap this study addresses is the absence of integrated accounts that connect these mechanisms within a single institutional strategy and examine their results together.

3. Materials and Methods

3.1. Setting

The study is situated at Iscte—University Institute of Lisbon, a public university in Portugal organized into several schools and distributed across two campuses, in Lisbon and in Sintra. The institution promotes the IN-Iscte project, whose central objectives are the reduction in dropout and the promotion of academic success, with particular attention to the first year and to students in vulnerable situations. The program is deliberately multidisciplinary in its governance: bringing together Academic Administration, Social Services, Soft Skills Lab, Pedagogical Council, Information Systems Development Service, and Quality Management; and mobilizing technical staff, peer mentors, class tutors, first-year coordinators, and course directors. This breadth constitutes a coordinated institutional effort, which is what allows detection to be coupled to response.

The methodological orientation is that of an institutional case study combining several strands of evidence: administrative and academic records analyzed through a rule-based alarm and through machine learning models, and descriptive evidence on the reach and reception of the AI literacy initiatives. The case study design is appropriate to the research questions, which concern how an integrated strategy operates and with what results in a real institutional setting, rather than the isolated effect of a single controlled intervention.

3.2. Data

All academic information at Iscte resides in the FenixEdu information system, which constitutes the principal data source for the detection components of the strategy. Two categories of variables are distinguished. Enrolment variables, which are available from the moment a student registers, include: gender; nationality and second nationality; country and city of residence; a derived "displaced" indicator; age; admission category and phase; enrolment date and mode; enrolment status; debt status; scholarship status; degree; school; campus; statutes (such as working student or special educational needs);

admission grade; and the placement option number indicating where the institution ranked among the student's preferences. Two of these fields, enrolment status and debt status, are administrative statuses that the information system stores at their current value rather than as of the registration date. They are used by the rule-based alarm described in Section 3.3 and are not part of the feature set supplied to the models described in Section 3.4. Academic variables, which become available at the end of the first semester, include for each curricular unit: enrolled and completed ECTS credits; the ratio of completed to enrolled credits; not-assessed (NA) credits; and the NA ratio. The detection components focus exclusively on first-time, first-year students of undergraduate programs, the population in which risk is most concentrated. The analyses draw on cohorts spanning the 2016/17 to 2024/25 academic years, with the precise window depending on the component.

The use of administrative data for intervention raises ethical considerations that the program treats as integral rather than incidental. These include the secure storage and encryption of data, the definition of who may handle nominally identified information and under what conditions, and the inherent tension between targeted intervention, which requires identification, and the protection of student privacy. Because the data are collected primarily for administrative management, their repurposing for risk detection must be carefully bounded, and the system is designed so that nominal identification serves only the purpose of timely, proportionate support.

Two distinct processing operations must be distinguished, because they rest on different legal bases and demand different obligations. The first is the operational early-warning system, which processes identifiable records in order to select students for an offer of support. This processing relies on the performance of a task carried out in the public interest under Article 6(1)(e) of the General Data Protection Regulation (Regulation (EU) 2016/679), namely the institution's statutory mission to monitor and promote academic success, rather than on consent. The second is the retrospective, statistical analysis reported in this article, which is a secondary use of the same records for scientific research and statistical purposes as provided for by Article 89(1) GDPR and Article 31 of Portuguese Law No. 58/2019 of 8 August. No data were collected from students for research purposes, and all figures reported here are aggregated and anonymized. Formal ethics committee approval was not required. Under Portuguese law, prior approval by an ethics committee is mandatory for clinical research on human beings as defined by Law No. 21/2014 of 16 April, whose material scope is limited to studies intended to discover or verify effects on health, health states or health-care interventions. That scope does not extend to educational or social science research conducted on anonymized institutional administrative records, and no equivalent requirement applies to research of the present kind. The safeguards required by Article 89(1) GDPR, in particular, data minimization and anonymization, were applied throughout. The operational system is governed by human oversight at every consequential step. The alarm produces a list for review rather than an automated decision, and no support, sanction or administrative consequence follows from a flag without a member of staff assessing the individual case. Contact takes the form of an open enquiry about how the student's year is progressing and whether any assistance would be useful; support is offered instead of imposed, and the decision to accept, partially accept, or decline it rests entirely with the student. Declining carries no recorded consequence, and flag status is not disclosed to the teaching staff responsible for assessing the student.

Students are informed at the start of the academic year about the range of support available to them, including the welcoming, mentoring and student support programs through which contact is subsequently made. They are not, however, told at the point of contact that an alarm has identified them, and the conversation is deliberately framed as

a general enquiry rather than as a response to a risk classification. The rationale is the stigmatization cost that may come from informing a student that an automated rule has categorized them as at-risk. This may itself damage the sense of belonging that the intervention exists to protect and may deter precisely those students least inclined to seek help. The cost of that choice is a reduction in transparency, and with it the absence of any procedure by which a student might contest a flag, since a student unaware of a classification cannot challenge it. What limits the significance of that absence is that a flag carries no adverse consequence of any kind: it produces an offer of assistance and nothing else, it triggers no administrative action, and it leaves no mark on the student's record.

Given the contextual nature of these contacts, the outcomes of the interactions they open cannot be specified in advance. They emerge instead as a co-constructed process, taking multiple forms depending on each student's profile, background and the specific situation at hand, and within that process the student plays an active role in the choices that shape their academic trajectory. Beyond this co-constructed nature, the establishment of contact can also be understood as an intermediate level of intervention, through which the student is informed about the range of responses and resources available within the university. In doing so, it translates, to some extent, the complexity of the institution and brings these responses closer to the student. Subsequent help-seeking, whether immediate or at a later stage, may therefore be considered an outcome of this contact.

3.3. Rule-Based Early Warning Alarm

The first detection component is a transparent, rule-based alarm that flags students at the end of the first semester according to three criteria, defined as follows:

- A student is flagged if the ratio of completed to enrolled ECTS credits is below 50%, indicating that the student completed less than half of the credit load in which they were enrolled;
- A scholarship-holding student is flagged if they completed fewer than 18 ECTS credits in the semester, a threshold that places the renewal of the scholarship at risk and therefore signals compounded academic and financial vulnerability;
- A student is flagged if their enrolment status has changed to a request for withdrawal (but not yet effective cancellation of enrolment).

These three categories are not mutually exclusive: a student may be flagged under more than one rule simultaneously. To support intervention, several enrolment variables are retained alongside the flag, so that the staff member making contact can access relevant contextual information and tailor the approach to the student's profile. Once flagged, students are reviewed case by case, and personalized intervention plans are established through negotiated dialogue between the student and the services assigned to follow up. The intervention infrastructure that the alarm feeds into includes a peer mentor team drawn from diverse academic backgrounds, health and well-being workshops, expanded transversal and soft skills training, pedagogical innovation training for staff, a dedicated "Student Space" bringing together several services, and a monthly routine of meetings among service representatives, student representatives, and teaching staff. The pilot phase of the early warning system analyzed data collected in March 2024 for the first semester of 2023/24 (Lopes et al., 2024); the impact of the resulting intervention was then assessed on the 2023/24 cohort, whose re-enrolment outcomes were observed at the start of 2024/25, set against the institution's longer-run trends in dropout and academic progress (Mauritti et al., 2025).

3.4. Machine Learning Early Detection

The second detection component investigates whether risk identification can be automated and brought forward in time using machine learning. For each academic year, two datasets were constructed: an enrolment-only dataset, available at the start of the year, and a dataset adding first-semester academic data, available at the end of the first semester. Both target two binary outcomes: academic success, defined as eligibility to progress to the second curricular year; and dropout, defined as the absence of enrolment in the following academic year. Preprocessing was standardized through a pipeline that one-hot encoded categorical variables, applied standard scaling to numerical variables, and imputed missing values as None for categorical fields and as −1 for numerical fields, ensuring that all datasets underwent identical treatment.

Seven widely used algorithms were trained and compared: Random Forest, Support Vector Machine (SVM), Naïve Bayes, Decision Tree, AdaBoost, K-Nearest Neighbors (KNN), and Logistic Regression. For every combination of model, dataset, and target, hyperparameter tuning was performed with the Optuna framework using 150 trials, and each model was then evaluated in its best-performing configuration. Four metrics were computed: accuracy, recall, F1-score, and the area under the receiver operating characteristic curve (ROC-AUC). Models were trained on data from 2016/17 to 2022/23 and tested on two datasets: a held-out subset of that same period, not used in fitting, and the 2023/24 cohort. The two-test design is deliberate with the goal of answering two different questions. The first establishes whether the models perform within the regime on which they were fitted (with no intention of offering evidence of generalization). The second asks whether that performance transfers to a cohort in which the institution had begun to intervene systematically. The 2023/24 cohort is distinctive because the personalized intervention strategy described above began that year; although the intervention itself was not included as a model variable, it may have altered student outcomes, allowing the comparison of pre- and post-intervention test performance to serve as a probe of model robustness under environmental change (Rodrigues et al., 2025).

Three features of this design limit what the results in Section 4.3 can support. First, the two test sets differ in kind and not merely in period. The pre-intervention test set is a random hold-out drawn from the pooled 2016/17–2022/23 data and therefore shares cohort year effects, admission grade scales, curricular structures and administrative coding conventions with the training data, whereas the 2023/24 test set is a genuine out-of-time evaluation. Any comparison between the two consequently confounds the change in period with the change in validation regime. Second, the features that were actually supplied to the models were audited at field level. The two administrative statuses that would most directly encode the dropout outcome, enrolment status and debt status, are stored by the information system at their extraction time value rather than as of the registration date, and enrolment status takes values including request for withdrawal and withdrawal; the audit established that neither field was part of the modelling feature set, so the most direct route by which a registration time feature could have carried the outcome is excluded. The audit does not extend beyond those fields: no time-stamped feature dictionary or observation time boundary was enforced for the remaining variables, and no feature ablation analysis was carried out, so residual leakage cannot be ruled out. Third, the reported recall and F1 values were computed with the majority class as the positive class, so they describe the identification of students who succeed rather than the detection of those at risk; class prevalence, per-split sample sizes and a trivial baseline were not recorded alongside them. Hyperparameters were tuned over 150 Optuna trials per configuration, with every preprocessing step fitted inside the training folds through scikit-learn pipelines, so that no test data entered the fitting of any transformation; the criterion by which the best trial was then selected was not recorded. For these reasons the machine learning

component is reported below as an exploratory analysis and as a methodological cautionary case, not as a validated detection system. The full pipeline, variable definitions and tuning procedure are documented in Rodrigues et al. (2025).

3.5. AI Literacy Initiatives

The third component, operating since May 2023, addresses the proactive equity mechanism through two complementary lines. The faculty line comprises master classes, pedagogical sessions, and a Specialization Seminar on GenAI in Teaching and Research. It is oriented toward a rigorous understanding of how generative models function, including their limitations and ethical implications, and toward the reconfiguration of assessment practices, privileging authentic assessment centered on critical analysis and the explicit articulation of reasoning over tasks easily delegated to AI systems. Rather than a logic of surveillance and fraud detection, the approach seeks to design academic activities whose very structure discourages uncritical use.

The student line is a curricular unit, "Academic Work with Artificial Intelligence", in operation since the 2023/24 academic year. It is mandatory in the first year of the ten undergraduate degrees on the Sintra campus and optional in the first year of the twenty undergraduate degrees on the Lisbon campus. Its content ranges from the basic functioning of the technology to the practical exploration of tools, the development of critical thinking through the analysis and comparison of AI-generated responses, the delimitation of academic integrity (distinguishing legitimate support, assisted co-authorship, and plagiarism) and the use of AI as a reasoning partner while preserving the central role of human judgment. The pedagogical rationale is that a universal first-year baseline guarantees a common level of AI literacy at the outset of the academic path, freeing later curricular units to focus on contextualized application within their specific fields. Evidence on these initiatives is necessarily more descriptive than the detection components and consists principally of measures of reach and participant reception. Because the unit is mandatory on one campus and optional on the other, it constitutes a universal baseline for entrants to the ten Sintra degrees and an opt-in offer for entrants to the twenty Lisbon degrees. The term *universal* is used below only in that campus-specific sense.

4. Results

4.1. Who Is at Risk? (RQ1)

The rule-based alarm provides a detailed map of how academic risk is distributed across the first-year population. In the 2023/24 pilot, approximately 19.2% of first-year students (n = 361) were identified as requiring specific intervention. The overwhelming majority of those flagged (98%, n = 354) had completed less than half of their enrolled credits. More troubling, 58% of flagged students (n = 209), corresponding to 11% of all first-year students, completed zero ECTS in the first semester, and approximately 22% of flagged students (n = 79), or 4% of all first-year students, had formalized a request for withdrawal. The close relationship between completing no credits and undertaking no assessment is striking: more than three-quarters of the zero-credit students have not been present for any assessment activity at all, indicating that the dominant pattern is not failure after effort but disengagement before it.

Risk is far from evenly distributed across student groups. Table 1 reports the incidence of the most severe indicator, completing zero credits (R0), across selected statuses and characteristics in the 2023/24 pilot. International students, and especially students from the Portuguese-speaking African countries (PALOP), stand out among the most exposed, with R0 rates of 36.1% and 42.9% respectively, against an institutional average of 11.1%. Relocated students (25.3%) and late-enrolment students (26.8%) are likewise

substantially above average, and the data on age reveal a clear gradient: while students under 20 have an R0 rate of 7.7%, the rate rises to 18.2% for those aged 20–29 and 34.0% for those aged 30–39, and stands at 23.6% among students aged 40 and above. These older students are concentrated in evening and post-work programs and frequently combine study with employment and family responsibilities, which plausibly accounts for their elevated risk.

Table 1. Incidence of severe academic risk (zero completed ECTS, R0) by student status and characteristics, first-year students, 2023/24 pilot cohort. Categories are neither mutually exclusive nor exhaustive, and the incidences are univariate and therefore not mutually adjusted. Wilson score 95% confidence intervals are reported for every cell above the disclosure threshold. Under the statistical disclosure control rule applied here, cells based on fewer than five students are reported as "n < 5" and their confidence intervals suppressed to avoid re-identification. For the same reason, the oldest age groups are combined into a single band of students aged 40 and above.

Student Status/Characteristic	N	R0 (%) [95% CI (%)]
Students from PALOP countries	91	42.9 [33.2–53.1]
International students	72	36.1 [26.0–47.6]
Late-enrolment students	71	26.8 [17.9–38.1]
Relocated students	249	25.3 [20.3–31.0]
Students with special educational needs	14	n < 5 [-]
Students in after-work programs	221	10.9 [7.4–15.6]
Working students	123	7.3 [3.9–13.3]
State grant scholarship holders	264	3.8 [2.1–6.8]
Students transferred to a new degree	42	n < 5 [-]
Class representatives	60	n < 5 [-]
Age < 20 years	1375	7.7 [6.4–9.2]
Age 20–29 years	406	18.2 [14.8–22.3]
Age 30–39 years	47	34 [22.2–48.3]
Age ≥ 40 years	55	23.6 [14.4–36.3]
All first-year students	1883	11.1 [9.8–12.6]

Equally informative are the apparent protective profiles. State grant scholarship holders are markedly below the institutional average at 3.8%, while class representatives and students who had transferred into a new degree show a lower incidence still, in both cases involving fewer than five students. These associations should, however, be read as artefacts of the way the categories are constituted rather than as protective effects. Class representatives are elected after the academic year has begun, so the category translates on a student being present and engaged at the time of election; and students transferred into a new degree have chosen to do so because the new degree more closely addresses their preferences, thus fueling engagement. Both are therefore colliders, and neither supports an inference that the status confers resilience. The position of scholarship holders is different. For a first-time first-year cohort, social action support is awarded on a means test at entry (or at least in the first weeks of the school year), and the progression thresholds that condition its renewal apply prospectively from the second year onwards. Scholarship status is therefore a prior characteristic for this cohort rather than an outcome, and the low incidence observed among these students is a genuine descriptive finding, albeit a univariate one that cannot on its own establish that the support confers protection. The counts are in any case too small to sustain a conclusion.

Very few students are affected in each of the three groups, and the resulting confidence intervals are so wide as to be uninformative. For students with special educational needs, the count falls below the disclosure threshold, so neither an incidence nor an interval is.

Table 1 is also univariate throughout, and its categories are neither exhaustive nor mutually exclusive. This matters most for the PALOP and late-enrolment rows: as noted in Section 2.2, visa delays mean that PALOP students frequently enroll late, so the elevated PALOP rate may be substantially or wholly mediated by late enrolment.

A conceptual distinction that the data force into view, and that is developed further below, is the one between academic failure and dropout. The alarm identifies students who are failing to progress, but failing to progress is not the same as leaving the institution. Keeping these two phenomena analytically separate is essential both for accurate institutional diagnosis and for the design of proportionate intervention.

4.2. Outcomes of Flagged Students by Intervention Status (RQ2)

Assessing whether detection coupled with intervention improves outcomes requires evidence that speaks to the intervention itself rather than to incidental differences between unrelated cohorts of first-year students. We therefore approach RQ2 at two levels. At the institutional level, we ask whether the indicators the strategy is designed to move (first-year dropout and academic progress) have shifted since personalized intervention began with the 2023/24 first-year cohort. At the individual level, and more decisively, we compare the subsequent outcomes of flagged students according to whether they actually received intervention, which provides the most direct available test of the intervention's contribution within a single cohort.

Table 2 places the two intervention years within the longer institutional series, combining two indicators measured on different populations. The first-year dropout rate (measured as the share of a given first-year cohort that fails to re-enroll the following academic year) fluctuated between roughly 14% and 21% across the cohorts from 2016/17 to 2022/23, with the highest values recorded in the cohorts most affected by the pandemic. For the 2023/24 and 2024/25 cohorts, the first two exposed to the personalized intervention strategy, this rate fell to 13.70% and 13.48% respectively, the two lowest values in the entire series. The average number of ECTS credits concluded tells a parallel story: it reached 44.29 in 2023/24 and 44.51 in 2024/25, the two highest figures observed since 2016/17, against a prior multi-year band of roughly 40 to 44. It is notable that both indicators reached their most favorable values in the two intervention years and continued to improve from the first to the second.

Table 2. First-year dropout and average concluded ECTS credits by academic year, 2016/17–2024/25. Dropout is defined as the proportion of the first-year cohort entering in that given year that fails to re-enroll at the start of the following academic year; the ECTS average is computed across all first-year enrolled students for the year indicated. Rows in bold are academic years in which the interventions took place. N is the number of students entering the first year in that cohort; the rise from 2021/22 to 2022/23 reflects the opening of a new School in Sintra, which increased the number of degrees offered and hence the number of places. The 95% confidence intervals reported for the dropout proportion are Wilson score intervals.

Academic Year	N	Average Concluded ECTS (First-Year Students)	First-Year Dropout at Start of Following Year (%) [95% CI (%)]
2016/17	1371	42.87	14.40 [12.6–16.3]
2017/18	1392	43.87	14.64 [12.9–16.6]
2018/19	1344	40.37	17.74 [15.8–19.8]
2019/20	1481	42.79	14.32 [12.6–16.2]

2020/21	1473	41.82	21.26 [19.2–23.4]
2021/22	1476	41.18	20.85 [18.9–23.0]
2022/23	1768	41.46	19.40 [17.6–21.3]
2023/24	**1883**	**44.29**	**13.70 [12.2–15.3]**
2024/25	**1907**	**44.51**	**13.48 [12.0–15.1]**

The comparison should not, however, be anchored on the pandemic peak. Measured against the four pre-pandemic cohorts of 2016/17–2019/20, whose mean first-year dropout was 15.28%, the improvement in the two intervention years (mean 13.59%) is about 1.7 percentage points, and measured against the lowest pre-pandemic year, 14.32%, it is about 0.7 points. The intervals in Table 2 are wide enough to consider the following: the interval for 2023/24 (12.2 to 15.3) overlaps that of every pre-pandemic cohort, and pooling the four pre-pandemic years (851 of 5588, 15.2%, 14.3 to 16.2) against the two intervention years (515 of 3790, 13.6%, 12.5 to 14.7) leaves intervals that overlap as well (even if on a very small margin). On its own, the series cannot distinguish the intervention years from the pre-pandemic baseline. Although these aggregate movements are consistent with the strategy working (the two most favorable years are the two intervention years rather than an isolated good year, which makes a one-off anomaly less likely), they cannot, however, isolate the strategy's effect on their own: observed at the level of whole cohorts, they remain open to alternative explanations and do not establish causation. It is therefore difficult to attribute the lower dropout values to the single reason of the intervention practice at the institution or, for that matter, to a pandemic recovery wave (which would be very difficult to measure), but it is worth considering that it has occurred against a macro-context (of economic disruption associated with the war in Ukraine) that would, if anything, be expected to worsen them. Therefore, Table 2 is presented as context only since the series cannot identify an effect, and no inference about the strategy is drawn from it here. For that reason, the analysis turns to a comparison internal to the 2023/24 cohort, which holds the surrounding environment fixed.

The most direct evidence comes from comparing, within the 2023/24 cohort, the re-enrolment outcomes of students according to both whether the early warning alarm flagged them and whether they subsequently received personalized intervention. Table 3 reports this breakdown across the four resulting groups. All 1883 first-year students of the 2023/24 cohort are included in the analysis that follows and all are observed at the start of 2024/25; no student was excluded at any stage, and no exclusion was made after re-enrolment status was known. Of the cohort, 361 students (19.2%) were flagged by the alarm and 1522 were not. Personalized intervention reached 45 of the 361 flagged students and a further 59 students who had not been flagged (these were students that voluntarily reached out to the Student Space and requested some kind of support and were given personalized interventions as well, depending on their context), giving 104 intervened students in all.

Table 3. Re-enrolment and dropout outcomes for first-year students of the 2023/24 cohort, by flag status and intervention status. Dropout is measured at the start of 2024/25. The four groups are mutually exclusive and exhaustive and together account for the full first-year cohort. Percentages are within-group percentages; 95% confidence intervals are Wilson score intervals.

Group	N	Re-Enrolled Following Year (%) [95% CI (%)]	Dropout (%) [95% CI (%)]
Flagged, intervened	45	68.9 [54.3–80.5]	31.1 [19.5–45.7]
Flagged, not intervened	316	43.7 [38.3–49.2]	56.3 [50.8–61.7]
Not flagged, intervened	59	93.2 [83.8–97.3]	6.8 [2.7–16.2]
Not flagged, not intervened	1463	95.8 [94.5–96.6]	4.2 [3.3–5.4]

Two findings stand out. First, the early warning system identifies genuine risk. Among students who received no intervention (the closest available approximation to an untreated baseline), flagged students dropped out at 56.3%, against only 4.2% for those not flagged, a more than thirteen-fold difference. Detection is therefore not generating noise: the students it singles out are, without support, those who overwhelmingly go on to leave. Second, the intervention is associated with a marked reduction in that risk. Among flagged students, those who received intervention dropped out at 31.1%, compared with 56.3% for those who did not, a risk difference of 25.2 percentage points (95% CI 10.6 to 39.8) and a relative reduction of 44.8% (95% CI 13.8 to 64.6; Fisher's exact p = 0.002). These intervals are wide because only 45 students received intervention: that group's rate of 31.1% carries a 95% interval of 19.5% to 45.7%, against 56.3% (50.8 to 61.7) among the 316 flagged students who did not. With groups of this size, the smallest difference detectable at 80% power and a two-sided 5% level is a fall from 56.3% to about 34%, some 22 percentage points; anything smaller could not have been distinguished from chance, however consequential in practice. By contrast, intervention made little difference among non-flagged students (6.8% versus 4.2%; risk difference 2.5 percentage points, 95% CI −4.0 to 9.0), exactly as expected for a group not exhibiting academic risk; the marginally higher rate among the few non-flagged students who were nonetheless intervened most likely reflects concerns not captured by the credit-based early warning system. Considered as a detector, the rule-based alarm can also be characterized at its operating point, which provides the natural benchmark for the models of Section 4.3. If we consider only the 1779 students who received no intervention, the closest available untreated sample, and taking the 240 dropouts among them as the positive class, the alarm attains a sensitivity of 74.2%, a specificity of 91.0%, a positive predictive value of 56.3% and a balanced accuracy of 82.6%, at an alert rate of 17.8%. These figures are reported here because three transparent, human-readable rules turn out to be a demanding baseline, a point returned to in Section 4.3.

It is important, however, to temper this reading with some caveats. First, intervention reached only a minority of flagged students, 45 of 361, or roughly one in eight. This is due to the fact that most of the flagged students, when contacted, declined the offer of help or support. Even though institutional capacity (as discussed in Section 2.3) could become a factor, this was not the case in this study because very few students ended up receiving some kind of intervention. Instead, this presents itself as an instance of the reluctance to seek help discussed in Section 2.1. Second, the comparison is observational rather than experimental: students who received intervention may differ systematically from those who did not, in reachability, engagement, or motivation, so part of the observed gap may reflect selection rather than the intervention alone. Third, students who do not re-enroll in their initial degree program but remain actively enrolled in another program at the University (i.e., internal program changes) are not classified as dropouts. However, the available data do not allow us to identify students who leave the University to continue their studies at another higher education institution. Consequently, some students classified as dropouts may in fact represent institutional transfers rather than withdrawal from higher education. Fourth, the data reaffirm that academic difficulty does not equate to departure: even among flagged students who received no intervention, 43.7% re-enrolled, underscoring that the flagged population remains substantially reachable and that failure and dropout must be kept analytically distinct. This last point is developed further below, since it conditions both how outcomes should be measured and where continued support is most likely to convert a difficult first year into eventual success.

Two further qualifications are more consequential than those above, and they materially limit what Table 3 can support. First, the composition of the flagged group is not independent of the outcome. Rule 3 flags students whose enrolment status has changed

to a request for withdrawal, and 79 students carried that status. Inspection of their records establishes the following. The rule adds no one to the flagged pool: every one of the 79 also failed a credit ratio criterion, 68 of them (86%) having completed no credits and the remaining 11 (14%) less than half of their enrolled credits, so all would have been flagged had Rule 3 never existed. Also, all 79 subsequently dropped out and none of them received intervention. The difficulty therefore lies not in the rule but in the status it records: a student who has already formalized a withdrawal request is near-certain to be recorded as a dropout and, having in effect already decided to leave, very unlikely to take up an offer of support. These students consequently occupy the non-intervened arm exclusively and, by representing a group with no realistic prospect of receiving the intervention, cannot contribute to a comparison between those who received it and those who did not. Setting the 79 aside and restricting attention to flagged students for whom intervention was a live possibility, dropout among the 237 remaining non-intervened students is 41.8% (99/237) against 31.1% among the 45 intervened, a risk difference of 10.7 percentage points with a 95% interval of −4.3 to 25.6 that includes zero (Fisher's exact $p = 0.19$). This contrasts with the unadjusted 25.2-point difference referenced above and, therefore, reported here for transparency. The 79 were retained in the operational cohort by design: it was not known, when the alarm was deployed, whether a student who had requested withdrawal could be brought back, so no attempt was made to triage them out and contact was offered to every flagged student on the same terms. No student was removed from the program, Table 3 reports the cohort in full, and the restricted comparison is set beside the unadjusted contrast rather than substituted for it. Second, no covariate adjustment has been attempted. The enrolment variables retained alongside the flag are the obvious adjustment set, and the account of help-seeking in Section 2.1 implies that willingness to accept support is patterned by exactly the vulnerabilities that also predict dropout, so the expected direction of the resulting bias is the direction observed.

The defensible reading of Table 3 is therefore that flagged students who engaged with support had better re-enrolment outcomes than those who did not, on a comparison this design cannot render causal. This cohort is too small and its intervened group too selected to settle the question either way.

4.3. Exploratory Predictive Modelling and Its Limits (RQ3)

The machine learning component addresses whether automated detection can scale beyond the transparent rule-based approach and generalize to new cohorts. Because the central question is how performance behaves when the environment changes, the full metric-by-metric results are reported rather than a summary. Table 4 reports the performance of the seven algorithms on the pre-intervention test set, drawn from 2016/17–2022/23, and Table 5 reports their performance on the 2023/24 cohort, the first exposed to personalized intervention. For each target (academic success and dropout) and each dataset (enrolment data available at the start of the year and the same data augmented with first-semester results) accuracy, recall, F1-score, and the area under the ROC curve (ROC-AUC, on a 0–100 scale) are shown. ROC-AUC is the most informative of the four metrics in this setting: because at-risk students are a minority, a model can post high accuracy and recall simply by favoring the majority class, whereas ROC-AUC reflects genuine discrimination between at-risk and not-at-risk students.

Table 4. Model performance on the 2016/17–2022/23 (pre-intervention) test set. The models were fitted on the 2016/17–2022/23 data and are evaluated here on a random hold-out drawn from that same pooled period, so these figures describe within-regime behavior and are not to be read as achievable out-of-sample performance; Table 5 reports the out-of-time evaluation. [a] Decision Tree: scores at or near the ceiling from registration-time data are reported as an artefact of the evaluation rather than as performance (see Section 4.3). [b] Degenerate rows: the model predicts virtually a single class, so accuracy, recall and F1 reproduce the prevalence of the majority class and carry no information about the detection of at-risk students.

Target	Dataset	Model	Accuracy	Recall	F1	ROC-AUC
Success	Start of the year	Random Forest	94	99	96	89
		SVM	79	98	87	62
		Naïve Bayes	74	86	83	63
		Decision Tree [a]	99	99	100	99
		AdaBoost	78	95	86	63
		KNN	82	95	88	71
		Logistic Regression	78	96	87	63
	End of the 1st semester	Random Forest	94	98	96	91
		SVM	95	99	96	91
		Naïve Bayes	89	92	82	86
		Decision Tree [a]	100	100	100	100
		AdaBoost	94	97	96	91
		KNN	94	99	96	89
		Logistic Regression	94	97	96	91
Dropout	Start of the year	Random Forest	97	100	98	90
		SVM [b]	85	100	92	56
		Naïve Bayes	80	90	88	59
		Decision Tree [a]	100	100	100	100
		AdaBoost [b]	85	99	92	57
		KNN	87	98	92	64
		Logistic Regression [b]	85	99	92	57
	End of the 1st semester	Random Forest	95	99	97	88
		SVM	94	99	97	85
		Naïve Bayes	89	92	93	84
		Decision Tree [a]	100	100	100	100
		AdaBoost	94	98	96	85
		KNN	94	99	96	85
		Logistic Regression	94	99	97	85

Table 5. Model performance on the 2023/24 (post-intervention) test set. A dash denotes an ROC-AUC that the analysis pipeline did not return. [b] Degenerate rows: the model predicts virtually a single class, so accuracy, recall and F1 reproduce the prevalence of the majority class and carry no information about the detection of at-risk students. [c] Logistic regression on the success target at the end of the first semester returned zero recall and F1; this is a fault in that fitting run rather than a property of the algorithm.

Target	Dataset	Model	Accuracy	Recall	F1	ROC-AUC
Success	Start of the year	Random Forest	79	97	88	58
		SVM	78	99	87	54
		Naïve Bayes	77	87	85	66
		Decision Tree	68	77	78	57
		AdaBoost	80	97	88	60

		KNN	77	91	86	59
		Logistic Regression	72	78	71	65
	End of the 1st semester	Random Forest	93	97	95	88
		SVM	93	98	95	86
		Naïve Bayes	87	86	91	88
		Decision Tree	71	69	79	75
		AdaBoost	89	99	93	76
		KNN	93	98	95	87
		Logistic Regression [c]	22	0	0	-
Dropout	Start of the year	Random Forest	85	96	92	56
		SVM [b]	86	100	92	-
		Naïve Bayes	80	88	88	59
		Decision Tree	71	77	82	55
		AdaBoost [b]	86	98	92	53
		KNN	84	95	91	54
		Logistic Regression [b]	86	100	92	-
	End of the 1st semester	Random Forest	89	92	93	80
		SVM [b]	86	100	92	-
		Naïve Bayes	78	78	86	79
		Decision Tree	86	90	92	76
		AdaBoost	91	98	85	72
		KNN	89	98	94	67
		Logistic Regression	91	98	95	71

The pre-intervention test set (Table 4) has to be read with two prior cautions, both related to rows that cannot be taken at face value. The first is the Decision Tree, which returns scores at or near the ceiling on every target and dataset, including an ROC-AUC of 100 from registration time data alone. That is not a credible level of performance on a task on which five of the remaining six algorithms do not exceed 71 from the same information. Therefore, the row is treated as an artefact of the evaluation instead of as a result. The second caution is the majority class convention noted in Section 3.4. A classifier that always predicts a majority class of 85% attains accuracy 85, recall 100 and F1 92, which is precisely the SVM row for start-of-year dropout, while AdaBoost and Logistic Regression report 85/99/92 on the same task. Those rows are degenerate, and their F1 values carry no information about at-risk detection. Precision for the at-risk class, the area under the precision–recall curve and a majority class baseline are the metrics a confirmatory evaluation would require.

Once those rows are set aside, Table 4 makes clear the contribution of academic data. Random Forest is consistently strong (ROC-AUC 88–91), including with enrolment data alone (89 for success, 90 for dropout). For the remaining five models, the value of academic data is weaker. With enrolment data only, their ROC-AUC clusters in a weak band (roughly 62–71 for success and 56–64 for dropout) even though accuracy and recall are frequently above 90 (a gap that exposes the class imbalance effect). Adding first-semester results lifts these same models to ROC-AUC values of roughly 84–91, confirming the well-established predictive power of early academic performance while also showing that enrolment-only prediction, though weaker, is feasible and therefore useful for intervening before results exist.

The post-intervention test set (Table 5) depicts a different story: the Decision Tree model, whose pre-intervention scores sat at the ceiling, has its ROC-AUC fall to 57 and 55 on the start-of-year success and dropout tasks, and it recovers only partially, to 75 and 76, even with first-semester data; the Random Forest model declines sharply on the start-of-

year tasks (to 58 for success and 56 for dropout), though it holds up better once academic data are included (88 and 80). More broadly, every model's start-of-year ROC-AUC sits in a near-random band of roughly 53–66, and several models degenerate entirely: SVM and Logistic Regression collapse into predicting a single class on the dropout target, and no ROC-AUC is reported for them because the analysis pipeline did not return one; a model that emits constant scores has an ROC-AUC of 0.5 rather than an undefined value, so the dashes in Table 5 record a gap in the analysis output and not a mathematical impossibility. Logistic Regression on the success target fails completely at the end of the first semester, with zero recall and F1. The models that retain the most value are the end-of-first-semester variants (ROC-AUC roughly 67–88), but even these fall below their pre-intervention levels.

One interpretation of this degradation would attribute it to model drift induced by the intervention itself: by supporting students who would otherwise have struggled, the strategy would have weakened the historical association between risk profiles and outcomes on which the models relied. However, the observed degradation should not be single-handedly attributed to intervention-induced model drift on the present evidence. There are three explanations that are more plausible. First, the pre-intervention figures were inflated by an artefact of the evaluation itself rather than by any property of the models. A Decision Tree scoring 100 on accuracy, recall, F1 and ROC-AUC from registration time data is not a result that a generalizing model could produce on this task, and target leakage is the standard explanation for performance of that kind (Kapoor & Narayanan, 2023). The field-level audit reported in Section 3.4 removes the most obvious candidate, since neither enrolment status nor debt status was among the features supplied to the models, but it does not establish what took their place: a residual field carrying post hoc administrative information and an unmodelled dependence between the fitting and hold-out partitions; both remain open and cannot be separated without re-fitting the pipeline. A model relying on a shortcut of either kind would perform near-perfectly in-period and collapse out of period, which is what Table 5 shows. Second, the validation design confounds the comparison. Pre-intervention performance was estimated using random hold-outs from pooled multi-year data, whereas 2023/24 performance was measured on a genuinely out-of-time cohort. Random-split validation commonly overstates temporal generalization, and the study did not perform the falsification test needed to distinguish this effect from intervention-induced drift. Third, the underlying data distribution and administrative process changed. Dropout prevalence fell from 17.5% across the training years to 13.7% in 2023/24, which could impair transfer independently of intervention effects. The defensible conclusion is therefore narrower: models selected on in-period random-split performance transferred poorly to the following cohort, with an unresolved artefact in the pre-intervention evaluation, the validation design, and distributional or administrative shift as plausible explanations that this study cannot separate. More generally, once institutions act on model predictions, historical risk models may require continuous out-of-time validation and recalibration (Lenert et al., 2019; Perdomo et al., 2020). On the present evidence, the interpretable three-rule alarm remains competitive, while the algorithmic detector has no demonstrated advantage.

4.4. Reach and Reception of the AI Literacy Provision (RQ4)

The AI literacy component addresses the proactive equity mechanism, and its results are best understood in terms of reach, design rationale, and reception rather than predictive performance. The faculty line has, since May 2023, delivered master classes, pedagogical sessions, and a Specialization Seminar on Generative AI in Teaching and Research that has run for eight editions and trained more than 220 faculty members. Participant reception has been uniformly positive, with all participants rating the training as relevant

or very relevant to their work. The orientation of this line toward authentic assessment, rather than detection and surveillance, reflects the recognition that the most durable safeguard against uncritical AI use is the design of tasks that reward reasoning that cannot be delegated.

The student line, the first-year course "Academic Work with Artificial Intelligence," has trained more than 2000 students since 2023/24. It is mandatory for entrants to the ten Sintra degrees and optional for entrants to the twenty Lisbon degrees, so it constitutes a universal baseline on one campus rather than across the institution. Its equity rationale is explicit: rather than leaving the acquisition of AI literacy to the uneven distribution of prior exposure and confidence, the institution guarantees a common baseline at the very start of the academic path, and it is the mandatory status of the unit on the Sintra campus, rather than the mere availability of the course, that constitutes the inclusion mechanism there. Because the course is mandatory across the ten Sintra degrees, every entering student on that campus acquires the same foundational understanding of how generative systems work and fail, how to evaluate the reliability of their outputs, how to distinguish legitimate support from plagiarism, and how to use AI as a reasoning partner rather than a substitute for judgment. The pedagogical literature underpinning this design holds that unsupervised use tends toward shallow, over-reliant engagement, whereas structured, teacher-mediated use can preserve and strengthen critical thinking (Heung & Chiu, 2025; Ramos et al., 2025). The teacher is thus positioned as an indispensable mediator, and the universal baseline frees later curricular units to concentrate on contextualized application within each field. While the available evidence is descriptive and does not yet permit causal claims about learning outcomes, the scale and reach of the initiative establish that a common literacy floor across an entering cohort is operationally achievable, which is the precondition for it to function as an equity measure. Whether the provision narrows a gap therefore remains untested. Establishing it would require a validated pre- and post-course AI literacy measure with subgroup breakdown, capable of demonstrating both baseline heterogeneity and post-course convergence, and preferably scenario-based reasoning tasks rather than a knowledge test alone; longer-run behavioral proxies such as academic integrity case rates would strengthen the argument further. Where the unit is optional, uptake is moreover exposed to precisely the self-selection dynamic identified in Section 2.1, under which an offer open to all is taken up disproportionately by those who already arrive confident. That is the mechanism the equity argument is intended to counter, and two-thirds of the institution's undergraduate degrees currently fall under the optional regime.

5. Discussion

Read together, the four strands of evidence describe a single equity strategy operating through two reinforcing mechanisms. The detection-and-support components constitute the reactive arm: because vulnerable students are unevenly distributed and frequently do not seek help, the institution identifies risk proactively and responds with differentiated, negotiated support; the within-cohort evidence indicates that flagged students who engaged with that support re-enrolled at higher rates than those who did not. The AI literacy components constitute the proactive arm: because an emerging competency is unequally distributed at entry, the institution builds a common baseline so that technological change does not widen existing gaps. The value of viewing these as facets of one strategy, rather than as separate projects, is that they share a single equity rationale and a single focus of action. Both rest on the premise that responsibility for the conditions of success lies with the institution rather than with the individual student, and both concentrate their effort on the first-year transition, where that responsibility is most consequential. An institution that pursued only the first (the reactive arm) would remain

perpetually downstream, treating difficulties after they arise; an institution that pursued only the second (the proactive arm) would build capacity while continuing to lose its most vulnerable students.

The distribution of risk documented in Section 4.1 is a direct argument against generic intervention. The students most exposed to severe early disengagement (PALOP and other international students, late-enrolment students, relocated students, and older students reconciling study with work and family) face qualitatively different obstacles, and the protective profiles show that status does not map onto outcome in any simple way. A support model that treated all flagged students identically would misallocate effort and miss the specific mechanisms of each group's difficulty. The program's design, in which flags trigger case-by-case review and negotiated, personalized plans drawing on a multidisciplinary infrastructure, is responsive to this heterogeneity. The within-cohort comparison underscores both the promise and the condition of the model: among flagged students, those who received intervention dropped out less often than those who did not, though not distinguishably so once students who had already formalized a withdrawal are set aside, and intervention reached only a minority of those flagged, constrained by students who, once contacted, declined support. Detection therefore produces benefit only when coupled to support that the student actually engages, which places a premium on outreach that is credible and accessible to precisely the groups least likely to seek help spontaneously.

The finding that academic failure does not automatically lead to dropout has consequences for both measurement and practice. For institutional metrics, conflating the two produces a distorted picture: a substantial share of students who fail to progress nonetheless re-enroll (43.7% of flagged students who received no intervention did so), and counting them as lost would both overstate dropout and obscure the population that remains reachable. For practice, the distinction defines an opportunity. Students who fail but persist are, by their persistence, signaling continued commitment, and they constitute a group for whom continued, adapted support may convert a difficult first year into eventual success. The strategy's separation of the two phenomena (failure and dropout) allows the institution to identify this group and to direct support accordingly rather than abandoning students whom a cruder metric would have written off. A caveat applies: because the data cannot distinguish genuine departure from transfer to another institution, some apparent dropout may in fact be continuation elsewhere, and confirming this requires direct contact with students.

Positioning AI literacy as an equity measure reframes a debate often conducted in terms of integrity and detection. If GenAI is treated primarily as a threat to be policed, the institutional response gravitates toward surveillance, which neither develops students' capabilities nor addresses the underlying inequality in who can use these tools well. If, instead, AI literacy is treated as a competency that all students have a right to acquire, the response gravitates toward universal provision and toward assessment redesigned to reward the reasoning that AI cannot supply. Where the unit is mandatory, as it is on the Sintra campus, it ensures that the capacity to use AI critically is not reserved for those who arrive already advantaged. Where it is optional, it does not: an offer open to all is precisely the arrangement under which prior confidence and cultural capital govern uptake. Extending the mandatory baseline to the Lisbon degrees is therefore the substantive equity step, and it has not yet been taken. The accompanying faculty development recognizes that this inclusion cannot be delivered by underprepared teachers, and the orientation toward authentic assessment aligns the incentive structure of coursework with the cultivation rather than the circumvention of critical thinking.

These findings can be read against the frameworks set out in Section 2, which the case corroborates in some respects and qualifies in others. Coulon's account of affiliation

to *le métier d'étudiant* predicts that a failure of affiliation will manifest early and behaviorally rather than as poor performance after sustained effort. The observation that more than three-quarters of zero-credit students sat no assessment at all is consistent with that prediction in a way that a grade-based account of first-year failure would not anticipate, and it suggests that the object of early intervention should be presence and engagement rather than academic remediation. The help-seeking literature predicts that those most in need are least likely to present themselves, and the finding that intervention reached roughly one in eight flagged students, limited by refusal after contact, is consistent with it. The same literature also implies that uptake is patterned by vulnerability, which is precisely why the observed outcome gap in Table 3 cannot be read as solely an intervention effect. The learning analytics literature anticipates that rule-based systems trade discovery for interpretability while machine learning systems offer earlier and potentially more accurate detection. The present case inverts that expectation: three transparent rules proved at least as dependable a detector, and the algorithmic models showed no demonstrated advantage over them. This suggests that in institutions of this size, the marginal return on algorithmic sophistication is smaller than the marginal return on intervention capacity, which is a finding about where scarce institutional effort should be directed. There is also a reason of principle, developed in Section 4.3, why a transparent rule should be expected to hold up better in this setting than a fitted model: it describes a student's present state rather than forecasting a future one, and so cannot be falsified by the institution acting to change that future. Finally, the competence framework literature treats digital and AI competence as conditioned by material prerequisites rather than by aptitude. The present study corroborates that a common baseline is deliverable at scale, but, lacking any pre- and post-course measurement, it cannot speak to whether that baseline narrows the gap the frameworks identify, and the optional regime on the larger campus stands in direct tension with the frameworks' implication that equal availability is not equal opportunity.

6. Limitations

Several limitations need to be considered. First, the study is a single-institution case, and although its purpose is to illuminate how an integrated strategy operates rather than to estimate a generalizable effect, the specific figures and dynamics reflect Iscte's context and should be transferred to other settings with care. Second, the definition of dropout as non-enrolment cannot distinguish departure from higher education altogether from transfer to another institution, so dropout figures may overstate genuine attrition. Third, the post-intervention observation window is short: the within-cohort comparison of intervention outcomes rests on a single cohort (2023/24), the institutional trend in dropout and academic progress covers only the two intervention years, and the machine learning analysis has only one post-intervention test year. This limits confidence in the durability of the observed improvements and in any interpretation of the degradation in predictive performance, and it means that the aggregate trend, however suggestive, cannot by itself be attributed to the strategy. Fourth, the evidence on the AI literacy mechanism is largely descriptive, establishing reach and reception but not learning outcomes, so the inclusion claim rests on design rationale and feasibility rather than on measured impact. Finally, the association between intervention and reduced dropout is observational rather than experimental: intervention was not randomly assigned and reached only a minority of flagged students, so those who received it may differ systematically from those who did not and the design does not isolate the program's causal contribution from this selection or from other concurrent changes. The very success of intervention in altering outcomes may be, as Section 4.3 suggested, simultaneously a confound for the predictive models. Further limitations bear directly on the strength of the claims made above. The within-cohort comparison rests on a restricted contrast, since the students who had formalized a

withdrawal request did not receive intervention, and no balance assessment or covariate adjustment has been performed. The predictive results are exploratory for the design reasons given in Section 3.4; the anomalous pre-intervention rows could not be diagnosed within the present analysis, which would require re-fitting the pipeline, and no calibration, at-capacity precision or subgroup evaluation was computed. The risk map is univariate and cannot mutually adjust the incidences; whether nationality and disability status should serve as model inputs at all, given that they proxy structural disadvantage but also risk encoding profiling into an automated triage applied to a vulnerable population, is a question this study raises without resolving.

7. Conclusions

This article has presented an integrated case study of an equity-oriented strategy at a Portuguese public university, arguing that securing equitable success for a diversifying student body requires both a reactive arm, which detects and supports those most at risk, and a proactive arm, which distributes emerging competencies inclusively, and that these are coherent facets of a single inclusion strategy concentrated on the first-year transition. The evidence indicates that academic risk is unevenly distributed, falling most heavily on international, PALOP, late-enrolment, relocated, and older students; that early disengagement, rather than failure after sustained effort, is the dominant pattern among the most severely affected; that academic failure does not automatically lead to dropout, leaving a reachable population of persisting students; that flagging identifies genuine risk, the three-rule alarm attaining a sensitivity of 74.2% and a positive predictive value of 56.3% at a 17.8% alert rate; that flagged students who engaged with support re-enrolled at higher rates than those who did not (considering limitations expressed in Section 4.2); that exploratory predictive models trained on pre-intervention cohorts transferred poorly to the following cohort, for reasons this design cannot separate from validation artefacts and an unresolved anomaly in the pre-intervention evaluation, and showed no demonstrated advantage over the transparent rule; and that a common first-year literacy baseline is operationally deliverable at scale on the campus where it is mandatory, which is a precondition for it to function as an inclusion measure rather than a demonstration that it does.

Several implications transfer beyond the specific institution. The coupling of detection to a multidisciplinary, negotiated response, rather than detection in isolation, appears fundamental, since flagging students without capacity to respond (or unwillingness to receive support) yields little results. The analytical separation of failure from dropout improves both diagnosis and the targeting of support and is straightforward for other institutions to adopt. The predictive component yields a prosaic but consequential caution. Administrative status fields that are updated over the course of an academic year can encode the outcome a model is meant to predict; in-period random-split validation flatters models that exploit them; and a transparent rule may match or exceed an algorithmic detector while remaining explicable to the staff who must act on it. Institutions deploying learning analytics on administrative data should specify the observation time of every feature, validate out of time, report performance for the at-risk class against a trivial baseline, and monitor subgroup performance before treating any model as operational. In addition, the framing of AI literacy as a universal entitlement instead of a matter to be policed offers a constructive posture amid a rapidly changing technological landscape.

Future work should pursue several directions. The predictive models would benefit from additional data sources not available in this study, including socio-demographic variables such as parental education and occupation, and engagement data from learning platforms such as Moodle, which may capture early disengagement before it appears in formal results. Longitudinal follow-up across more cohorts would clarify whether the lower dropout and improved academic progress persist and would help separate the

program's causal contribution from the selection inherent in who takes up support and from broader environmental change. The AI literacy mechanism, in particular, requires evaluation that moves beyond reach and reception to measured effects on critical thinking, academic integrity, and equitable outcomes.

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

Funding: This work was financially supported by Project IN-Iscte - Espaço para Crescer II, Call 05/C06-i07/2023, funded by the Portuguese Recovery and Resilience Program (PRR), The Portuguese Republic and The European Union (EU) under the framework of Next Generation EU Program. The work has also been partially supported by Portuguese national funds through Fundação para a Ciência e a Tecnologia, I.P. (FCT) under project UID/04466/2025 (DOI: https://doi.org/10.54499/UID/04466/2025).

Institutional Review Board Statement: Ethical review and approval were waived for this study. Under Portuguese law, mandatory prior approval by an Ethics Committee (and, in particular, by a Health Ethics Committee or the National Commission for Clinical Research Ethics) is required for clinical research on human beings, as defined and regulated by Law No. 21/2014 of 16 April (the Clinical Research Act). The material scope of that Act (Article 1) is limited to studies intended to discover or verify effects on health, health states or health-care interventions; it does not extend to educational or social-science research. Portuguese law does not impose a requirement of prior Ethics Committee approval for studies of this kind that rely on anonymised institutional administrative records. Accordingly, formal Institutional Review Board / Ethics Committee approval was not required for this study. For completeness, the study complies with the principles for processing personal data for scientific-research and statistical purposes set out in Article 89(1) of the General Data Protection Regulation (Regulation (EU) 2016/679, GDPR) and in Article 31 of Law No. 58/2019 of 8 August (which implements the GDPR in Portugal), namely data minimisation and anonymisation of the data.

Informed Consent Statement: Written informed consent from participants was not required and no consent form exists. No data were collected directly from students for the purposes of this research; the study constitutes a secondary, statistical use of administrative records lawfully held by the institution. Where such records contain personal data, their processing relies on the performance of a task carried out in the public interest (Article 6(1)(e) GDPR) rather than on consent, and the further processing for scientific-research and statistical purposes is expressly permitted, subject to anonymisation, under Article 89(1) GDPR and Article 31 of Law No. 58/2019. Recital 26 GDPR further confirms that the data-protection principles do not apply to anonymous information, including where processed for statistical or research purposes. As all data reported are anonymised and aggregated, no individual is identifiable and the requirement of informed consent for research participation does not apply.

Data Availability Statement: The data presented in this study are available on request from the corresponding author due to privacy restrictions associated with personally identifiable data.

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

Abbreviations

The following abbreviations are used in this manuscript:

AI	Artificial Intelligence
ECTS	European Credit Transfer and Accumulation System
PALOP	Portuguese-Speaking African Countries ("Países Africanos de Língua Oficial Portuguesa")
NA	Not Assessed
RQ1-RQ4	Research Question 1–4
R0	Zero Completed ECTS
SVM	Support Vector Machine
KNN	K-Nearest Neighbors
ROC-AUC	Area Under the Receiver Operating Characteristic Curve
F1	F1-Score
PRR	Portuguese Recovery and Resilience Plan ("Plano de Recuperação e Resiliência")
EU	European Union

References

Almeida, L. S., Guisande, M. A., & Paisana, J. (2012). Extra-curricular involvement, academic adjustment and achievement in higher education: A study of Portuguese students. *Anales de Psicología*, *28*(3), 860–865.

Baker, R. S., & Hawn, A. (2022). Algorithmic bias in education. *International Journal of Artificial Intelligence in Education*, *32*(4), 1052–1092. https://doi.org/10.1007/s40593-021-00285-9.

Behr, A., Giese, M., Teguim Kamdjou, H. D., & Theune, K. (2020). Early prediction of university dropouts—A random forest approach. *Jahrbücher Für Nationalökonomie Und Statistik*, *240*, 743–789. https://doi.org/10.1515/jbnst-2019-0006.

Bozkurt, A. (2024). Why generative AI literacy, why now and why it matters in the educational landscape? Kings, queens and GenAI dragons. *Open Praxis*, *16*(3), 291–306. https://doi.org/10.55982/openpraxis.16.3.739.

Capucha, L., Pereira, C., & Godinho, R. (2021). *Avaliação do contributo do PT2020 para a promoção do sucesso educativo, redução do abandono escolar precoce e empregabilidade dos jovens.* Available online: https://portugal2020.pt/wp-content/uploads/relatorio_final_contributo_pt2020_educacao.pdf (accessed 13 April 2026)

Casanova, J. R., Bernardo, A. B., & Almeida, L. S. (2021). Difficulties in academic adaptation and intention to drop out of students in the first-year of Higher Education. *Revista de Estudios e Investigación En Psicología y Educación*, *8*(2), 211–228.

Cosgrove, J., & Cachia, R. (2025). *DigComp 3.0: European digital competence framework for living, working and learning.* Publications Office of the European Union.

Costa, A. F., & Lopes, J. T. (2011). The diverse pathways of higher education students: A sociological analysis on inequality, context and agency. *Portuguese Journal of Social Science*, *10*(1), 43–58.

Costa, A. F., Lopes, J. T., & Caetano, A. (Eds.). (2014). *Percursos de estudantes no ensino superior: Fatores e processos de sucesso e insucesso*. Mundos Sociais.

Costa, E. B., Fonseca, B., Santana, M. A., de Araújo, F. F., & Rego, J. (2017). Evaluating the effectiveness of educational data mining techniques for early prediction of students' academic failure in introductory programming courses. *Computers in Human Behavior*, *73*, 247–256. https://doi.org/10.1016/j.chb.2017.01.047.

Coulon, A. (2017). Le métier d'étudiant: L'entrée dans la vie universitaire. *Educação e Pesquisa*, *43*(4), 1239–1250.

European Union. (2024). *Regulation (EU) 2024/1689 of the European parliament and of the council of 13 June 2024* (Artificial Intelligence Act). European Union.

Gardner, J., Brooks, C., & Baker, R. (2019). Evaluating the fairness of predictive student models through slicing analysis. In *Proceedings of the 9th international conference on learning analytics & knowledge* (pp. 225–234). Association for Computing Machinery. https://doi.org/10.1145/3303772.3303791.

Gil, P. D., Martins, S. da C., Moro, S., & Costa, J. M. (2021). A data-driven approach to predict first-year students' academic success in higher education institutions. *Education and Information Technologies*, *26*(2), 2165–2190. https://doi.org/10.1007/s10639-020-10346-6.

Hellas, A., Ihantola, P., Petersen, A., Ajanovski, V. V., Gutica, M., Hynninen, T., Knutas, A., Leinonen, J., Messom, C., & Liao, S. N. (2018). Predicting academic performance: A systematic literature review. In *Proceedings companion of the 23rd annual ACM conference on innovation and technology in computer science education (ITiCSE 2018)* (pp. 175–199). Association for Computing Machinery. https://doi.org/10.1145/3293881.3295783.

Heung, Y. M. E., & Chiu, T. K. F. (2025). How ChatGPT impacts student engagement from a systematic review and meta-analysis study. *Computers and Education: Artificial Intelligence*, *8*, 100361. https://doi.org/10.1016/j.caeai.2025.100361.

Kapoor, S., & Narayanan, A. (2023). Leakage and the reproducibility crisis in machine-learning-based science. *Patterns*, *4*(9), 100804. https://doi.org/10.1016/j.patter.2023.100804.

Lenert, M. C., Matheny, M. E., & Walsh, C. G. (2019). Prognostic models will be victims of their own success, unless…. *Journal of the American Medical Informatics Association*, *26*(12), 1645–1650. https://doi.org/10.1093/jamia/ocz145.

Lopes, A. L., Mauritti, R., Martins, S. d. C., Pintassilgo, S., Matias, C., & José, S. (2024). Digital tools for the prevention of dropout and academic failure: A case study of a Portuguese university. In *EDULEARN24 proceedings* (pp. 7092–7098). IATED. https://doi.org/10.21125/edulearn.2024.1678.

Mah, D.-K., & Groß, N. (2024). Artificial intelligence in higher education: Exploring faculty use, self-efficacy, distinct profiles, and professional development needs. *International Journal of Educational Technology in Higher Education*, *21*, 58. https://doi.org/10.1186/s41239-024-00490-1.

Martins, S. d. C., Carvalho, H., Ávila, P., & Costa, A. F. (2017). Policies for widening participation and success factors in Portuguese higher education. *Creative Education*, *8*(2), 210–230.

Martins, S. d. C., Mauritti, R., & Machado, B. (2023). How has the pandemic affected access and the feeling of belonging in Portuguese higher education? *Education Sciences*, *13*(9), 942. https://doi.org/10.3390/educsci13090942.

Massano-Cardoso, I. M., de Carvalho Figueiredo, S., & Galhardo, A. (2024). Ansiedade, depressão e stress em estudantes universitários deslocados da sua residência. *Revista Portuguesa de Investigação Comportamental e Social*, *10*(2), 1–15.

Matias, A. R., Pinto, P. F., Rodrigues, V., Martins, P., & Seabra, T. (Eds.). (2023). *Trovoada de ideias: Inclusão linguística e social dos estudantes internacionais dos PALOP no ensino superior português*. OM/ACM.

Mauritti, R., Pintassilgo, S., Belchior-Rocha, H., Palma Saleiro, S., & Mendes, J. (2023). Mobilidade internacional por motivos de estudo: Fluxos e distribuição de estudantes da CPLP no ensino superior e território português. *Cidades, Mobilidades e Territórios*, *47*, 176–192.

Mauritti, R., Roque Ferreira, M., Pintassilgo, S., & Lopes, A. (2025). Enhancing student success and reducing dropout: A case study from a Portuguese university. In *EDULEARN25 proceedings* (pp. 6484–6489). IATED. https://doi.org/10.21125/edulearn.2025.1594.

Oliveira, A., Mauritti, R., & Pereira, E. R. (2021). Políticas de democratização do acesso e diversificação do perfil de estudantes no Ensino Superior: Experiências de Portugal e Brasil. *Revista FORGES*, *7*(1), 82–101.

Pargman, T. C., & McGrath, C. (2021). Mapping the ethics of learning analytics in higher education: A systematic literature review of empirical research. *Journal of Learning Analytics*, *8*(2), 123–139. https://doi.org/10.18608/JLA.2021.1.

Perdomo, J., Zrnic, T., Mendler-Dünner, C., & Hardt, M. (2020). Performative prediction. In *Proceedings of the 37th international conference on machine learning*. PMLR.

Ramos, I., Mesquita, A., Lopes, A. L., Moreira, F., Barros, V., Sousa, R. D., Regadas, G., Gomes, L., & Leite, P. (2025). Studying the GenAI's impact on learning and critical thinking: A multi-site field experiment. In *ICIS 2025 proceedings*. AIS eLibrary.

Realinho, V., Machado, J., Baptista, L., & Martins, M. V. (2022). Predicting student dropout and academic success. *Data*, *7*(11), 146. https://doi.org/10.3390/data7110146.

Rodrigues, D. M., Lopes, A., Mauritti, R., Roque Ferreira, M., & Pintassilgo, S. (2025). Early detection of at-risk university students using machine learning: A study of model performance in evolving academic environments. In *EDULEARN25 proceedings* (pp. 6538–6545). IATED. https://doi.org/10.21125/edulearn.2025.1610.

Tlili, A., Shehata, B., Adarkwah, M. A., Bozkurt, A., Hickey, D. T., Huang, R., & Agyemang, B. (2023). What if the devil is my guardian angel: ChatGPT as a case study of using chatbots in education. *Smart Learning Environments*, *10*, 15. https://doi.org/10.1186/s40561-023-00237-x.

UNESCO. (2024). *AI competency framework for students*. UNESCO.

Vuorikari, R., Kluzer, S., & Punie, Y. (2022). *DigComp 2.2: The digital competence framework for citizens*. Available online: https://publications.jrc.ec.europa.eu/repository/handle/JRC128415 (accessed on 19 August 2026).

White, M., & Canning, E. A. (2023). Examining active help-seeking behavior in first-generation college students. *Social Psychology of Education*, *26*, 1369–1390. https://doi.org/10.1007/s11218-023-09794-y.

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