Revista Brasileira de Saúde

ISSN 3085-8089 vol. 2, n. 1, 2026

ARTIGO 12
Data de Aceite: 06/01/2026

INTEGRACIÓN DE IA PREDICTIVA Y BLOCKCHAIN PARA PRONOSTICAR EL COMPORTAMIENTO ESTUDIANTIL EN LA EDUCACIÓN SUPERIOR: UN MARCO CONCEPTUAL Y UN ANÁLISIS MULTICASO

Raúl Jaime Maestre
Associate Professor at ESIC Business & Marketing School, Barcelona, Spain
Doctorando en Ciencias de la Computación (UNIR); tesis sobre aplicación de algoritmos blockchain a la gestión de certificados educativos (prev. 2027). Formación de posgrado en IA/blockchain e interacción con dirección financiera; Licenciado en ADE y Diplomado en Ciencias Empresariales. Experto en transferencia tecnológica: consultor estratégico en proyectos de blockchain y fintech desde 2012, con trayectoria previa en administración y finanzas.
Profesor PDI en ESIC en ESIC Marketing and Businss School y docencia universitaria en UOC (TFM en blockchain y ciberseguridad) y en escuelas de negocio (EBS, IEBS y EALDE), enfocada en proyectos ágiles, industria 4.0 y transformación digital.
https://orcid.org/0000-0003-3669-8992

Abstract: Universities need to anticipate students' decisions (enrollment, switching, progression, dropout). Learning analytics and credential governance have been treated as separate domains. This creates a gap: AI generates insights, but a secure, auditable, institution-wide decision layer to operationalize them is missing. We pose three RQs: (how to combine academic, financial, and engagement data to predict behavior), (the role of blockchain in ensuring integrity, auditability, and governance), and (the organizational capabilities required to deploy an integrated analytics layer). We adopt a design-oriented, multi-case approach and propose the UCAS architecture, which integrates prediction with blockchain-based governance and credentials. We analyze six institutions using public documents and comparative thematic coding. Three findings emerge: first, predictive AI exists but in silos, without cross-unit orchestration; second, blockchain is used for credential issuance and verification, not as a governance layer for the behavioral data lifecycle; third, integration occurs when predictions are coupled to traceable operational triggers. Contribution: a model and roadmap to personalize services, improve retention, and align sustainability with privacy and traceability.

Keywords: Predictive algorithms; blockchain technology; consumer behavior; universities; artificial intelligence; institutional management; transactions

Introduction

Educational institutions face the challenge of understanding and adapting to student behavior, as they are considered their primary consumers (Alsaadi & Bamasoud, 2021). Factors such as program selection, learning preferences, and financial transactions require universities to adopt a strategic approach to meet students' needs and expectations (Baker & Inventado, 2014). In this context, predictive algorithms enable the analysis of large volumes of data to forecast student trends and behaviors, facilitating informed decision-making (Syed Mustapha, 2023).

On the other hand, blockchain technology ensures the integrity and transparency of digital interactions, such as the management of academic credentials, payments, and other administrative processes (Nakamoto, 2009). The integration of both technologies represents a significant opportunity to optimize institutional management and enhance the educational experience of students (Casino et al., 2019).

The recent literature on learning analytics and educational data science has largely treated artificial intelligence as a predictive instrument for modeling individual student performance, dropout risk, program completion likelihood, or engagement patterns. In parallel, research on blockchain in higher education has primarily focused on credential issuance and verification, academic record portability, and the traceability of learning achievements. These two streams – AI-driven predictive analytics and blockchain-enabled credential governance – have evolved mostly in isolation. Critically, there is very limited work that connects student behavioral analytics – treating the student not only as a learner but also as a decision-making consumer of institutional services – with blockchain-based data governance infrastructures capable of operationalizing those predictions in a coordinated, auditable, institution-wide way. This lack of integration represents a structural gap: universities are increasingly able to forecast what students are likely to do, but they lack a trusted mechanism to translate those forecasts into consistent academic, financial, and administrative action.

This is not only a technical limitation; it is a strategic one. The ability to anticipate enrollment decisions, switching across academic pathways, financial stress, or potential withdrawal has direct consequences for (i) student retention and, by extension, institutional revenue sustainability; (ii) academic portfolio planning and the allocation of teaching and support resources; and (iii) administrative efficiency in areas such as financial aid, targeted advising, and the recognition and communication of student achievements. Under growing competition, rising operating costs, and student expectations of personalized service, the absence of an integrated decision layer that links behavioral prediction, data traceability, and accountable institutional response constrains universities' capacity to manage both academic success and the economic relationship with each student.

This paper addresses that gap through a dual contribution. First, it proposes an integrated conceptual model – here referred to as the University Consumer Analytics Stack (UCAS) – which articulates three layers: (1) multi-source data collection and modeling of academic performance, financial behavior, and interactional signals from students; (2) AI-driven behavioral forecasting aimed at estimating continuation likelihood, enrollment intent, payment risk, and support needs; and (3) a blockchain-based governance and orchestration layer that enables traceability, auditability, and institution-wide execution of targeted interventions. Second, the paper applies and refines this model through a multi-case analysis of higher education institutions that have already begun to implement, although often in a fragmented manner, components of these layers. In doing so, the goal is not only to describe current practice, but to offer university leadership a roadmap for converting advanced analytics into reproducible, transparent, and strategically aligned decision-making at scale.

Literature Review. Current situation of artificial Intelligence algorithms in university consumer behavior

Predictive algorithms rely on machine learning techniques to identify patterns in large datasets (Casino et al., 2019). Their application in the university context includes:

1. Enrollment prediction: Predictive models analyze historical data to forecast enrollment in various academic programs. Factors such as labor market trends and previous preferences help design marketing strategies and academic planning (Rani, Sachan & Kukreja, 2023).
2. Personalization of services: Algorithms analyze students' preferences in using resources such as libraries or learning platforms to provide personalized recommendations that enhance their educational experience (Sharples & Domingue, 2016).
3. Academic performance management: By identifying patterns associated with academic success or failure, early interventions such

DOI https://doi.org/10.22533/at.ed.8208126020112

as personalized tutoring can be implemented to improve student performance (Tapscott & Tapscott, 2016).

Research questions

This study is guided by three core research questions that address both the analytical and organizational requirements for deploying AI- and blockchain-enabled decision intelligence in higher education.

RQ1. How can academic performance data and transactional/engagement data about the student-as-consumer be combined to predict enrollment intention, continuation likelihood, and dropout risk at the individual level?

This question targets the predictive layer. It assumes that student behavior is not only academic (grades, progression, attendance) but also economic and service-oriented (tuition payment patterns, program switching behavior, platform usage, interaction with support services). The goal is to understand whether these heterogeneous data streams can be fused to generate reliable early-warning signals that are actionable for the institution.

RQ2. What role can blockchain play in ensuring the integrity, traceability, and governance of those behavioral predictions across academic, financial, and administrative units within the university?

Here the focus is not on prediction accuracy, but on trust and accountability. If AI models are generating individual-level forecasts (e.g., "high dropout risk" or "likely to defer payment"), under what technical and governance conditions can those forecasts be shared, audited, and acted upon without creating opaque decision-making or regulatory exposure (e.g. GDPR non-compliance, bias claims)? This question positions blockchain not only as a credential layer, but as a coordination and accountability layer for institutional decision-making.

RQ3. What organizational capabilities must a university develop to operationalize an integrated AI–blockchain analytics stack in practice (e.g., processes, roles, governance structures, intervention protocols)?

This question addresses institutional readiness. It assumes that predictive insight alone does not create impact unless the university can embed it into workflows: targeted advising, adaptive financial aid, personalized academic pathways, automated alerts, etc. We ask which capabilities (data governance, cross-unit coordination, ethical oversight, intervention design, student communication) are required to move from pilots to sustained, institution-wide deployment.

Optional hypotheses for quantitative validation.

- H1a. Models that fuse academic with financial and engagement features achieve higher accuracy than academic-only models.
- H1b. Real-time engagement features enable earlier detection of high-risk students than lagging academic indicators.
- H2. A blockchain based audit layer increases perceived procedural fairness and accountability among internal stakeholders.
- H3. Cross-unit data governance is positively associated with the rate of interventions executed after alerts.

DOI https://doi.org/10.22533/at.ed.82082126020112

Conceptual framework: The University consumer analytics stack (UCAS)

This study proposes the University Consumer Analytics Stack (UCAS), an integrated conceptual model that connects (i) multi-source student data, (ii) AI-driven behavioral prediction, and (iii) blockchain-based governance and execution. The purpose of UCAS is to articulate how universities can move from fragmented analytics initiatives toward an institution-wide decision layer that is both predictive and accountable. The model positions the student simultaneously as a learner and as a service consumer whose academic, financial, and engagement behaviors have strategic implications for retention, resource allocation, program design, and institutional sustainability.

UCAS is organized into three interdependent layers: (1) the Data Layer, (2) the Predictive Intelligence Layer, and (3) the Governance and Orchestration Layer. Together, these layers define how data is captured, transformed into foresight, and converted into coordinated institutional action.

Data Layer: multidimensional student behavioral signals

The first layer of UCAS consolidates heterogeneous data streams associated with each student into a unified analytical view. We group these streams into three categories:

1. Academic performance and progression data: This includes grades, assessment histories, course completion status, attendance patterns, remediation needs, use of tutoring resources, and indicators of academic stress or underperformance. These variables represent the traditional domain of learning analytics, which seeks to model academic success and risk.
2. Financial and administrative transaction data: This includes tuition payment timing, scholarship and aid utilization, outstanding balances, deferral requests, installment behavior, and interactions with financial services. From a management and sustainability perspective, these variables are critical because they capture the economic dimension of the student–institution relationship. They also often provide early signals of instability, such as financial distress that precedes withdrawal.
3. Engagement and interaction data: This includes platform usage (LMS logins, time-on-task, submission punctuality), advising requests, career services interactions, programme change inquiries, attendance at onboarding or retention-related interventions, and even micro-signals such as reduced responsiveness to official communications. These variables reflect how actively the student is participating in the broader service environment provided by the university.

Conceptually, UCAS treats these three categories not as separate operational silos (academic affairs, finance, student services), but as facets of a single behavioral profile. The underlying assumption is that student continuation, program loyalty, perceived value, and dropout risk are multi-factor decisions that cannot be inferred from aca-

DOI https://doi.org/10.22533/at.ed.82082126020112

demic metrics alone. In other words, progression is both an academic trajectory and a consumption decision.

Within the UCAS model, the output of this layer is an integrated, temporally ordered student activity record that can be used to construct predictive features. This requires data interoperability and, in practice, some degree of shared schema across departments.

Predictive intelligence layer: Forecasting student decisions

The second layer uses artificial intelligence and advanced analytics to transform multi-source behavioral data into forward--looking institutional signals. This layer performs three core predictive functions:

1. Enrollment and continuation forecasting: Models estimate the likelihood that an admitted or currently enrolled student will (re) enroll in the next academic period, switch programs, or exit the institution. This includes identifying "at-risk" students before they formally disengage and detecting segments likely to respond to specific forms of support (academic advising, flexible payment plans, alternative course sequencing, etc.).
2. Dropout and disengagement risk assessment: By combining academic underperformance indicators with financial strain signals and declining engagement patterns, this layer estimates short-term withdrawal probability. Importantly, in UCAS the relevant output is not only "high risk / low risk," but also the inferred drivers of that risk (academic overload, unmet financial need, weak programme fit, etc.), because different drivers imply different interventions.
3. Personalization and pathway optimization: Predictive models recommend tailored actions at the individual level: alternative course pathways, modular or micro-credential options, financial counseling, contact from a retention advisor, or proactive eligibility review for aid. This is where analytics shifts from diagnostic ("who is at risk?") to prescriptive ("what should we offer, to whom, and when?").

The Predictive Intelligence Layer is not only technical. It encodes institutional intent. Which outcomes are we trying to optimize — graduation rate, perceived student satisfaction, lifetime tuition revenue, regulatory compliance, equity of access? UCAS makes this explicit: prediction is a governance instrument, not just a data science exercise.

The output of this layer is a set of actionable predictions and recommended interventions, each tied to specific students or cohorts, with associated confidence levels and rationales.

Governance and orchestration layer: Blockchain-enabled accountability

The third layer governs how predictions are stored, shared, audited, and activated across the university. UCAS frames this layer as blockchain-enabled for three reasons:

1. Data integrity and provenance: Storing hashes or state commitments of key analytic outputs (e.g., “student X flagged for financial risk on [date] due to [driver]”) on a tamper-evident ledger creates an auditable trail. This allows the institution to demonstrate that interventions were data-driven, applied consistently, and not arbitrary or discriminatory. It also protects against post hoc manipulation of records for compliance or reputational reasons.
2. Cross-unit coordination and access control: Academic services, financial aid, and student success offices often operate on different systems and under different regulatory constraints. Smart contract–mediated access policies can define who is allowed to view or act on a given alert, under what conditions, and with what obligations to log the response. In UCAS, blockchain acts as a governance substrate that enforces shared rules across organizational silos.
3. Intervention orchestration: UCAS conceptualizes certain interventions (tuition plan adjustments, advisor outreach, eligibility review for micro-credentials or modular pathways) as triggerable events. Smart contracts can be used to register these events, require acknowledgment by responsible units, timestamp actions taken, and record outcomes. This creates both traceability (“we acted within 48 hours of the risk alert”) and the possibility of measuring intervention effectiveness over time.

In this sense, the Governance and Orchestration Layer is not just about credential issuance — the dominant use case in current blockchain-and-education literature — but about institutional accountability in behavioral decision-making. It translates predictive insight into auditable, policy-aligned action.

Flow of insight to action in UCAS

UCAS is designed as a pipeline:

1. Data layer to predictive intelligence layer: The university continuously aggregates academic, financial, and engagement data and constructs behavioral profiles.
2. Predictive intelligence layer to governance and orchestration Layer: AI models generate forecasts (e.g., “Student A: 78% probability of non-continuation next term, likely driven by cumulative payment delays + low LMS activity”). These forecasts are written, referenced, or anchored in a controlled ledger environment that documents both the alert and its justification.
3. Governance and orchestration layer to institutional intervention: Based on pre-defined policies encoded as smart contracts or workflow rules, the relevant unit (financial aid office, academic advisor, success coach, program director) is prompted to intervene. The intervention is logged, along with timing and type (e.g., “offered revised installment schedule” / “adaptive study plan proposed”).

DOI https://doi.org/10.22533/at.ed.82082126020112

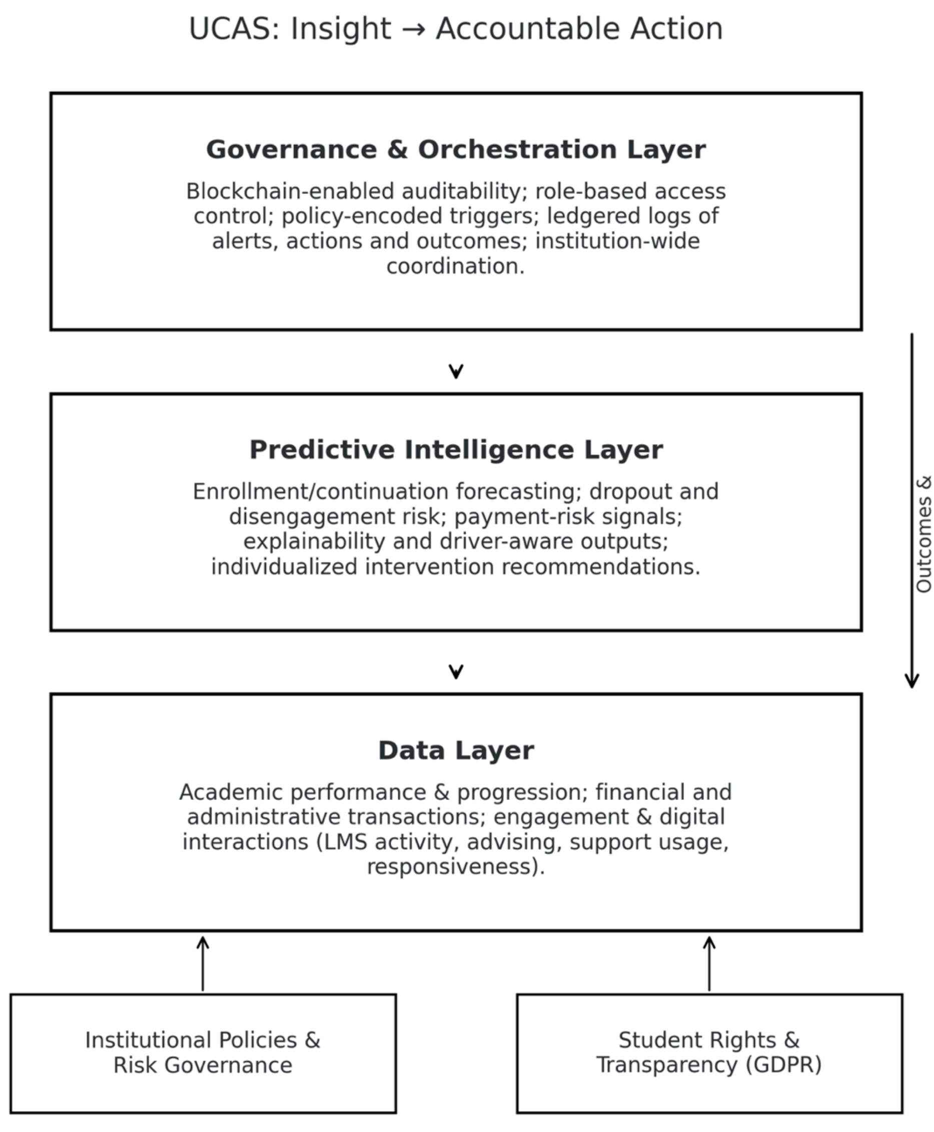

Image 1. UCAS: Insight -> Accountable Action

4. Feedback loop: The outcomes of interventions (student continues? resolves balance? re-engages with coursework?) are captured and feed back into the Data Layer, improving future model calibration and informing policy refinement.

This cyclical flow is central to UCAS, prediction without intervention is inert; intervention without auditability is opaque; auditability without multi-source data is superficial. The model therefore treats data fusion, behavioral forecasting, and blockchain-governed execution as mutually dependent capabilities rather than isolated projects.

Theoretical Positioning

Conceptually, UCAS contributes to three strands of theory:

- Learning analytics and student success research: UCAS reframes risk analytics from a purely academic construct ("likelihood of failing a course") to a behavioral-economic construct ("likelihood of discontinuing the institutional relationship"), aligning student success with institutional sustainability.
- Data governance and trust in AI-mediated decision making: By embedding auditability and access control into the architecture, UCAS positions blockchain as more than a credential registry. It frames it as an organizational trust mechanism that legitimizes the use of algorithmic predictions in high-stakes decisions affecting students.
- Strategic management of higher education: UCAS conceptualizes the university as an integrated decision system, where academic quality, financial stability, and service personalization are co-optimized through predictive analytics. This moves beyond the traditional separation between "student affairs," "academic affairs," and "finance," and instead models the student as a shared object of strategic action.

Propositions

To guide empirical examination, the UCAS framework implies the following high-level propositions:

- P1. Universities that integrate academic, financial, and engagement data into a unified behavioral profile will generate more accurate and earlier risk forecasts than institutions relying on academic signals alone.
- P2. The presence of a blockchain-governed orchestration layer increases internal accountability and cross-unit coordination in student interventions, thereby raising the likelihood that predictive insights are translated into timely, concrete institutional actions.
- P3. The effectiveness of UCAS is contingent on organizational readiness: without defined intervention workflows and governance structures, even high-quality predictions will not materially improve retention, progression, or perceived student value.

These propositions connect the theoretical model to evaluation. They also create a bridge into the following section (Methodology), where UCAS can be examined empirically (through multi-case analysis) or validated structurally (through a design science approach).

Methodology

This study adopts a design-oriented, multi-case qualitative methodology. The objective is twofold: (i) to develop and formalize the University Consumer Analytics Stack (UCAS) as an integrated decision architecture for higher education, and (ii) to examine to what extent elements of this architecture already exist in leading institutions, and how they are being used to inform academic, financial, and service decisions about students.

The methodology follows a logic consistent with Design Science Research (DSR) in information systems — in which an artifact (here, UCAS) is proposed to address an identified organizational problem — combined with comparative multi-case analysis to assess the contextual relevance, feasibility, and institutional preconditions for that artifact to function in practice.

The section is structured as follows: Section 3.1 describes the research design; Section 3.2 explains case selection, data sources, and data collection procedures; Section 3.3 outlines the analysis strategy; Section 3.4 discusses validity, reliability, and limitations; and Section 3.5 summarizes ethical and governance considerations relevant to working with student-related analytics.

Research design

The research design is anchored in the three Research Questions (RQ1–RQ3):

- RQ1: How can academic, financial, and engagement data about the student-as-consumer be combined to predict enrollment intention, continuation likelihood, and dropout risk?
- RQ2: What role can blockchain play in ensuring integrity, traceability, and governance of those predictions across academic, financial, and administrative units?
- RQ3: What organizational capabilities are required for universities to operationalize such an integrated AI–blockchain analytics stack?

To address these questions, we proceed in two stages:

Artifact development (conceptual design stage). We synthesize prior literature on learning analytics, student retention analytics, financial risk monitoring in higher education, and blockchain-based credential governance to construct UCAS, an integrated three-layer model (data, predictive intelligence, governance/orchestration). This corresponds to the classical DSR sequence of problem definition to objective specification to artifact design.

Contextual evaluation (multi-case stage). We then analyze multiple higher education institutions that have implemented advanced analytics, blockchain-enabled data governance, or both. The objective is not to "test" UCAS in a statistical sense, but to (i) map which layers of UCAS are already present and how they are instantiated; (ii) identify gaps that prevent full institutional

DOI https://doi.org/10.22533/at.ed.82082126020112

integration; and (iii) infer the organizational capabilities that enable or block operationalization. This evaluation step corresponds to DSR's relevance cycle and environment grounding.

This combined approach is appropriate for an area where full, production-level integration of AI-driven behavioral forecasting with blockchain-governed orchestration is still emergent. In such contexts, explanatory case logic is more informative than controlled experimentation, because it reveals institutional constraints, governance tensions, and translation frictions between prediction and action — all of which are central to our research questions.

Case selection and data collection

Case selection logic

We use theoretical sampling rather than random sampling. Institutions were selected because they exhibit one or more of the following characteristics:

1. Mature learning analytics / predictive student success systems. The institution uses advanced analytics to forecast academic performance, progression, or dropout at the individual student level.
2. Operational use of behavioral or transactional data for decision-making. The institution incorporates financial, engagement, or service-interaction data (e.g., tuition payment behavior, advising interactions, LMS usage) into its student risk models, retention strategies, or resource allocation.
3. Adoption of blockchain for credentialing, data integrity, or auditability. The institution issues verifiable academic credentials onchain or uses distributed ledger infrastructure to ensure traceability of records and/or to support compliance and accountability in student-related decision processes.
4. Evidence of institutional coordination across academic, financial, and administrative units. The institution has structures (committees, integrated student success offices, shared governance protocols, automated workflows) that transform analytic insight into targeted intervention.

Universities that meet at least two of these criteria are considered high-value cases because they illustrate either partial instantiations of UCAS layers or credible organizational pathways toward full integration.

The resulting sample includes large research-intensive universities, specialized digital or online universities, and institutions publicly recognized for early adoption of blockchain-enabled credentialing. This diversity is intentional: it allows comparison across governance models (centralized vs. federated), student populations (traditional vs. non-traditional / lifelong learners), and strategic priorities (academic excellence, retention/revenue stability, credential portability, regulatory signaling).

Data sources

Data for each case was drawn from multiple convergent sources, including:

- Public institutional strategy documents, digital transformation roadmaps, accreditation self-reports, and student success initiatives.

DOI https://doi.org/10.22533/at.ed.8208212602011 2

- Technical descriptions of analytics platforms, retention dashboards, enrollment forecasting systems, and blockchain credentialing infrastructures (when disclosed by the institution or its technology partners).
- Policy documents and governance frameworks describing how student data, predictive alerts, and intervention decisions are shared across academic affairs, student success units, and financial services.
- Speeches, interviews, statements, and conference material from institutional leadership (e.g., Provost, CIO, VP for Student Success, VP for Enrollment Management), in which they explicitly describe analytics-driven decision processes.
- Regulatory and compliance communications (e.g., references to privacy, bias, auditability, external accreditation requirements, and traceability of decisions affecting students).

All documents were collected as publicly available materials or internal-facing-but-publicly-cited summaries (for instance, presentations at sector conferences that outline analytics architecture or blockchain credential pilots). No personally identifiable student data or confidential student records were accessed. The unit of analysis is institutional practice and governance, not individual students.

Data Recording

For each institution, we constructed a structured case profile capturing:

- Which UCAS layers are present (Data, Predictive Intelligence, Governance/Orchestration).
- Which data types are explicitly integrated for decision-making (academic, financial, engagement/service).
- Which predictive outcomes are targeted (enrollment forecasting, dropout risk, payment default, pathway personalization).
- Whether blockchain or other ledger-like infrastructures are used and for what purpose (credential issuance, auditability, access control, workflow orchestration).
- How interventions are triggered, delivered, and tracked.
- What organizational entities are responsible for acting on those triggers.

These profiles serve as the basis for cross-case comparison in Results.

Analysis strategy

The analysis proceeded in three steps.

Step 1: Layer mapping

For each case, we coded the evidence to determine which components of UCAS were present. For example, if an institution uses machine learning to generate individualized dropout-risk alerts that are routed to academic advisors, that maps to the Predictive Intelligence Layer plus partial orchestration. If an institution issues blockchain-verifiable diplomas but does not use behavioral analytics to trigger interventions, that maps primarily to the Governance and Orchestration Layer in its credentialing form.

Step 2: Capability extraction

We then identified the organizational capabilities that appeared necessary to make each layer operational. These capabilities include data integration infrastructure, cross-unit governance bodies, intervention protocols, compliance and ethics review mechanisms, and mechanisms for accountability (e.g., audit logs, service-level expectations for advisor outreach). The goal of this step is to answer RQ3: Which capabilities are preconditions for deploying UCAS in practice?

Step 3: Cross-case comparison and proposition refinement

Finally, we compared cases along three axes aligned with the theoretical propositions:

- Depth of data fusion (Are academic, financial, and engagement signals unified at student level?).
- Degree of intervention orchestration (Do predictions lead to coordinated, trackable actions, or are they dashboards with no operational follow-through?).
- Level of auditability and accountability (Are interventions and decisions traceable, contestable, and governed under shared policy, or are they ad hoc and opaque?).

This comparative analysis allowed us to refine the UCAS propositions, especially P1 (integration improves predictive value), P2 (blockchain-enabled governance improves accountability and intervention execution), and P3 (organizational readiness is a determinant of real impact). Where cases showed divergence from UCAS expectations, those divergences were used to qualify the model and identify constraints.

Validity, reliability, and limitations

Construct validity

We explicitly defined each UCAS layer and pre-specified what counted as evidence of that layer. This reduces subjective interpretation when coding cases (e.g., "predictive analytics" is not any dashboard; it must generate forward-looking, student-level risk or intent estimates and inform decisions).

Internal validity

Causal claims are treated cautiously. We do not infer that implementing UCAS automatically "causes" improved retention or financial stability. Instead, we examine whether institutions that approximate UCAS report that predictive insights are embedded into repeatable interventions, and whether governance infrastructures exist to hold units accountable for acting on those insights. Causal framing is therefore institutional and procedural, not statistical.

Reliability

All case profiles were created using the same template and coding logic. When similar language (e.g., "personalized learning," "student success analytics") was used by different institutions to describe different realities, we recoded it according to UCAS definitions rather than institutional marketing language.

External validity / generalizability

The cases are not claimed to be representative of the global higher education sector. They are theoretically informative (extreme or leading-edge cases) rather than statistically sampled. The findings therefore speak to feasibility, architectural patter-

DOI https://doi.org/10.22533/at.ed.82082126020112

ns, and governance requirements for early adopters. They provide propositions that future quantitative work can test at scale (e.g., linking specific data-integration capabilities to observed retention outcomes).

Limitations

The study relies on institutional selfreporting, strategic documents, and public technical descriptions. We do not directly audit internal databases, measure model accuracy in situ, or observe live student interventions. As a result, some claims made by institutions (e.g., on "real-time personalization" or "automated early warning") may reflect aspirational or pilot-stage capability rather than full operational maturity. We treat these claims analytically but do not assume their full realized impact unless supporting procedural evidence is documented.

Ethical, privacy, and governance considerations

Because UCAS explicitly connects student-level behavioral predictions to institutional action, ethical and regulatory considerations are inseparable from methodological rigor. Three principles guided both model construction and case analysis:

Data minimization and proportionality

The integration of academic, financial, and engagement data raises the risk of intrusive profiling or discriminatory targeting (e.g., treating students from lower-income backgrounds as "financial risk" cases). Any operationalization of UCAS must include explicit proportionality tests: Is the data being used necessary and justifiable for the stated intervention?

Transparency and contestability

If predictive labels (e.g., "likely to drop out") trigger differential treatment, students must be able to understand and contest those classifications. We therefore pay particular attention to whether institutions use blockchain or audit trails to log how and why an intervention decision was made, and who had access to the triggering information.

Compliance and bias governance

Predictive analytics in education can reproduce structural bias (e.g., penalizing non-traditional learners, working students, or students with caregiving responsibilities). The Governance and Orchestration Layer in UCAS is intentionally specified not only as a ledger, but as a mechanism for monitoring procedural fairness. In coding each case, we examine whether any governance body is explicitly responsible for fairness, equity, or non-discrimination in the use of predictive analytics.

In summary, the methodology combines design science (to articulate UCAS as an actionable institutional architecture) with comparative case analysis (to evaluate how close real institutions are to that architecture, and what capabilities are required to operationalize it). This approach directly supports the research questions by linking data integration, predictive intelligence, and blockchain-enabled governance to concrete patterns of retention strategy, resource allocation, and student-facing intervention.

Design and development of a university enrollment prediction algorithm

Predicting the number of enrollments is essential for a university's strategic planning. This includes resource allocation, quota definition, academic program offerings, and budgeting. A system based on predictive algorithms enables informed decision-making by utilizing historical data to identify patterns and trends (Rani, Sachan & Kukreja, 2023).

Data collection and preparation

The following sources of information can be used to train the algorithm:

- Historical enrollment data (Casino et al., 2019).
- Demographic characteristics of applicants, such as age, gender, and geographic location (Sharples & Domingue, 2016).
- Previous academic records, including grade point averages and exam scores (Rani, Sachan & Kukreja, 2023).
- External sources, including socioeconomic indicators and labor market trends (Tapscott & Tapscott, 2016).

It is important to consider how the data preparation will be carried out:

1. Data cleaning: Remove null, duplicate, or inconsistent values to improve data quality (Sharples & Domingue, 2016).
2. Normalization: Scale numerical variables to ensure all have equal weight during model training (Casino et al., 2019).
3. Categorical variable encoding: Convert non-numeric data (such as gender or region) into a format that the algorithm can process, such as One-Hot Encoding (Rani, Sachan & Kukreja, 2023).

Finally, selecting the relevant features is essential, as not all data are useful. The features that contribute most to predicting enrollments should be chosen:

- Student's academic history (Casino et al., 2019).
- Interaction with digital marketing campaigns, such as email clicks and time spent on the website (Sharples & Domingue, 2016).
- Academic programs of interest (Rani, Sachan & Kukreja, 2023).
- Socioeconomic level and distance from the university campus (Tapscott & Tapscott, 2016).

This can be done using techniques such as variable correlation analysis or feature importance analysis with models like decision trees (Casino et al., 2019).

Suggested artificial Intelligence algorithms for the design and development of the prediction algorithm

Depending on the available data, different machine learning models can be used for prediction. Below are the main options that can be applied.

Multiple linear regression

This model is suitable for a simple approach that explains how independent variables affect the number of enrollments (Montgomery et al., 2012). The base equation is as follows:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_n x_n$$

Where:

- y is the number of enrollments.
- $x_1,x_2,\ldots,x_n$ are the predictor variables.
- β_0 is the intercept.
- $\beta_1,\beta_2,\ldots,\beta_n$ are the coefficients associated with each variable.

The advantages of using multiple linear regression include its ease of interpretation for the algorithm user and its ability to identify linear relationships between different variables (James et al., 2013). On the other hand, a disadvantage is that this model cannot capture complex patterns in the data (Hastie et al., 2009).

Random Forest

A more advanced model than multiple linear regression, which combines several decision trees to make predictions (Breiman, 2001). It is useful for capturing non-linear relationships between variables.

The process that can be used would be:

1. Create multiple decision trees from different subsets of the dataset.
2. Average the predictions of the trees to obtain the final result.

The advantages of using random forests include their ability to handle complex and non-linear relationships in the data, as well as their ability to automatically identify the most important variables in the model (Liaw & Wiener, 2002). On the other hand, a disadvantage is that this model can be slower and more costly in terms of computational implementation (Fernández-Delgado et al., 2014).

Neuronal networks

This type of model is especially suitable for problems with large volumes of data and highly non-linear relationships (LeCun, Bengio & Hinton, 2015). Neural networks use multiple layers of interconnected neurons to learn complex patterns in the data.

The suggested architecture would be as follows:

- Input: Predictor variables such as age, region, or academic interest.
- Hidden layers: 2-3 hidden layers with activation functions like ReLU (Rectified Linear Unit) (Nair & Hinton, 2010).
- Output: A single neuron that returns a continuous value to predict the number of enrollments.

Neural networks are more accurate than other models due to their ability to analyze complex data and generalize non-obvious patterns (Schmidhuber, 2014). Another important advantage is their flexibility to capture non-linear relationships between variables. However, they have disadvantages, such as the need for a large and clean dataset to train the model adequately (Goodfellow, Bengio & Courville, 2016). Additionally, their interpretation is more difficult compared to simpler models.

Implementation hypothesis

División de datos Data Division

It is proposed to divide the data into three subsets to maximize the efficiency of model training (Goodfellow, Bengio & Courville, 2016):

- Training (70%): To train the model.

- Validation (15%): To tune hyperparameters.
- Testing (15%): To evaluate the final performance.

Evaluation metrics

The following metrics will be used to evaluate the model's performance (Chollet, 2017; Goodfellow et al., 2016):

- Mean absolute error (MAE):

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i|$$

Image 2. Mean absolute error

- Root mean squarre error (RMSE):

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}$$

Image 3. Root mean squarre error

- Coefficient of Determination (R2): Measures how well the model explains the variance of the data.

Practical example of using a university enrollment prediction algorithm

Let's suppose a university wants to predict enrollments for the next semester using an algorithm.

1. Available data:
   - Number of students who have applied to the university in the last 5 years.
   - Conversion rates (percentage of accepted students who enroll).
   - Advertising campaigns conducted.
   - Socioeconomic factors of the region.

2. Model used: The chosen model is Random Forest due to its ability to handle complex nonlinear relationships and work with a large set of heterogeneous data (Breiman, 2001).

2. Prediction: The algorithm predicts that, under current conditions, a 5% increase in enrollments is expected due to the success of a recent advertising campaign targeting international students. This prediction provides valuable information for strategic decision-making, such as resource allocation and academic planning (Chollet, 2017).

Development of the University services personalization algorithm

Personalizing services is key to improving the student experience at universities (Chen et al., 2020). This process involves adapting services such as academic programs, wellness resources, extracurricular activities, and communication strategies based on students' preferences and behaviors (Xu, 2024). A predictive algorithm allows for the analysis of large volumes of data to anticipate students' specific needs, increasing their satisfaction and improving overall performance (Almalawi et al., 2024).

Data collection and preparation process

The main sources of information that can be used for the university services personalization algorithm include:

- Academic data: Enrolled programs, academic performance, attendance level (Kurni, Mohammed & Srinivasa, 2023).
- Demographic Data: Age, gender, place of residence, socioeconomic level (Munir et al., 2023).
- Digital behavior: Interactions on the educational platform (LMS), website searches, email clicks (Ngulube & Ncube, 2025).
- Declared preferences: Surveys, registration forms, selected extracurricular activities (Kazem et al., 2022).
- Service usage: Participation in services such as tutoring, scholarships, libraries, sports activities (Gkontzis et al., 2019).

To prepare the data, the following steps should be followed (Maestre et al., 2023):

- Data Cleaning: Remove duplicates, outliers, and incomplete data.
- Normalization and Scaling: Convert all variables to a uniform range to prevent some features from dominating the model.
- Categorical Variable Encoding: Transform qualitative variables (such as faculty or gender) into numerical formats using methods like One-Hot Encoding.
- Label Creation: Group students according to similar needs, such as high academic risk students or highly engaged students (Zhang et al., 2021).

Selection of relevant features

To generate this model, it is crucial to identify the most important features, which may include:

- Frequency of digital platform use: Evaluate the time and interaction on educational platforms (Ngulube & Ncube, 2025).
- Grade history: Analyze previous academic performance to predict support needs (Kurni, Mohammed & Srinivasa, 2023).
- Extracurricular activities: Consider the level of participation in activities outside the classroom to identify additional interests and needs (Kazem et al., 2022).
- External factors: Place of residence and socioeconomic level as predictors of access to certain resources (Munir et al., 2023).

To use these features, the following techniques can be applied (Maestre et al., 2023):

- Correlation analysis: evaluate the relationship between variables and the desired outcome.
- Feature importance analysis: Use models like Random Forests to measure the relevance of each variable (Gkontzis et al., 2019).

Suggested predictive AI algorithms for university services personalization

Various approaches can be used depending on the data and analysis objectives (Kazem et al., 2022). Below are the main models suitable for generating the algorithm.

Model 1: Student Segmentation

K-Means is a clustering algorithm that divides students into different groups based on common characteristics (Maestre et al., 2023).

The process can be carried out as follows:

1. Select the number of groups (k) using a method like the elbow method (cluster analysis). This method is used to determine the number of clusters in a dataset (Gkontzis et al., 2019).
2. Assign each student to the nearest group based on their distance to the centroids.
3. Iteratively readjust the centroids until the internal variance of the group is minimized (Kurni, Mohammed & Srinivasa, 2023).

To apply the process as indicated:

- Segment students according to profiles: academically lagging students, students interested in specific extracurricular activities, etc. (Ngulube & Ncube, 2025).

The advantages of using this method are that it is easy to interpret and quick to implement, making it ideal for creating personalized service groups. However, it is sensitive to noisy data and variable scaling.

Model 2: Decision trees and Random forests

Decision trees are useful for predicting specific services that should be recommended to a student based on their characteristics (Zhang et al., 2021). Random forests combine multiple trees to improve accuracy.

The process can be carried out as follows:

1. Train the tree with data on student characteristics and their past choices.
2. Evaluate the importance of the features to identify which variables most influence personalization.
3. Use the random forest to reduce the risk of overfitting and improve generalization (Almalawi et al., 2024).

To apply the process as indicated:

- Predict which services (tutoring, scholarships, activities) are most relevant to a student.

The advantages of using this method are that it is robust against incomplete or noisy data and provides clear explanations of which features influence the recommendation (Xu, 2024).

Model 3: Neural Networks

Suitable for complex personalizations based on large volumes of data, such as students' digital interactions (Ngulube & Ncube, 2025).

The architecture can be designed as follows:

- Input layer: Student features (demographics, service usage, digital interactions).
- Hidden layers: Two or three layers with ReLU or Sigmoid activations.
- Output layer: Vector with specific recommendations (e.g., probability of using tutoring services, inter-

est in sports activities, etc.) (Kurni, Mohammed & Srinivasa, 2023).

To apply the process as indicated:

- Predict multiple needs simultaneously (e.g., probability of enrolling in specific courses and participating in sports activities) (Munir et al., 2023).

The advantages of using this architecture are that it is excellent for complex and nonlinear data, and it can learn interactions between multiple variables. However, this architecture requires large amounts of data and a longer training time for the predictive algorithm (Gkontzis et al., 2019).

Evaluation of the model used

The metrics to evaluate the performance of the predictive algorithm for university services personalization include:

- Accuracy: How well the model correctly classifies or predicts (Brownlee, 2021).
- F1 Score: Balance between precision and recall (Kazem et al., 2022).
- Confusion Matrix: Evaluate specific errors in predictions (Zhang et al., 2021).
- Mean squared error (MSE): In regression models, measures the difference between predicted and actual values (Maestre et al., 2023).

Implementation of the personalization algorithm

We should follow these steps for implementation, as determined by the process:

1. Integration of the model into an LMS (Learning Management System): Display personalized recommendations in real-time, such as learning resources or events (Chen et al., 2020).
2. Personalized notifications: Send emails or messages based on the algorithm's predictions (e.g., invitations to activities that align with the student's interests) (Kurni, Mohammed & Srinivasa, 2023).
3. Visualization dashboard: Provide reports to administrators on the general and specific needs of students (Almalawi et al., 2024).

Practical example of using a university services personalization prediction

A university implements a system based on random forests to personalize its services (Gkontzis et al., 2019).

- Input: Demographic data, previously selected activities, and LMS usage.
- Prediction: The model indicates that 70% of first-year students need math tutoring and 30% are interested in marketing activities (Kazem et al., 2022).
- Action: The university organizes tutoring sessions and promotes marketing activities among these groups.

Development of the academic perfomance management algorithm

Students' academic performance is one of the key factors in the success of universities. Predicting academic performance allows institutions to identify at-risk students, personalize intervention strategies, and optimize resources to improve educational outcomes (Maestre et al., 2023). This algorithm is designed to anticipate students' performance based on historical, demographic, and behavioral data, using machine learning techniques and advanced analytics (Gkontzis et al., 2019).

Data collection and preparation process

The main sources of information that can be used for the academic performance management algorithm include:

- Academic data: Partial and final grades from previous courses, number of courses taken and passed, and cumulative weighted average (CWA) (Kurni, Mohammed & Srinivasa, 2023).
- Demographic data: Age, gender, socioeconomic level, geographic location (Munir et al., 2023).
- Learning behavior: Use of educational platforms (LMS), participation in tutoring, class attendance (Ngulube & Ncube, 2025).
- External factors: Internet connectivity level, part-time employment, family environment (Xu, 2024).

To prepare the data, the following steps should be followed (Maestre et al., 2023):

- Data cleaning: Remove duplicates, incomplete, or inconsistent records.
- Variable transformation: Encode categorical variables (e.g., gender) using techniques like One-Hot Encoding.
- Normalization: Scale numerical variables to a uniform range to prevent any from dominating the calculations.
- Label generation: Classify performance as "high," "medium," or "low" based on predefined thresholds (e.g., CWA > 4.0 = "high performance").

Selection of relevant features

To select the features, the following techniques can be used:

- Correlation analysis: Identify the variables most related to academic performance (e.g., class attendance, previous grades) (Gevorgyan, 2025).
- Feature importance: Use models like Random Forests or XGBoost to measure the influence of each variable on the predictions (Breiman, 2001).

Key features typically include:

- Previous cumulative average.
- Participation in extracurricular activities.
- Use of LMS platforms (number of logins, resources downloaded).
- Socioeconomic factors (e.g., family income).

Design of predictive algorithms for academic performance management

Various approaches can be used depending on the data and analysis objectives. Below are the main models suitable for generating the algorithm.

Model 1: Logistic regression for classification

A statistical model that predicts the probability of a student belonging to a category (e.g., high performance).

The formula that can be used for this logistic regression model for classification would be the following:

$$P(y=1|X)=\frac{1}{1+e^{-(\beta_0+\beta_1X_1+\beta_2X_2+\ldots+\beta_nX_n)}}$$

Image 4. Regression model for classification

The application would be to classify students according to their probability of achieving "high" performance in their studies.

The advantages of using this model are that it is simple to implement and interpret, and it is ideal for binary or multiclass problems (e.g., high, medium, or low performance) (Hosmer et al., 2013).

Model 2: Random forests

Random forests combine multiple decision trees to improve accuracy and avoid overfitting (Breiman, 2001).

The process to follow would be:

1. Train multiple trees on random subsets of the data.
2. Combine their predictions by voting to classify or averaging for regression.

The application would be to identify at-risk students and predict the most influential variables in performance (Cutler et al., 2007).

The advantages of using this model are that it is robust against missing or noisy data and can identify the relative importance of features.

Modelo 3: Neural networks

The neural network model is an advanced model capable of identifying complex nonlinear relationships in the data.

The architecture can be designed as follows:

- Input layer: Student features (previous grades, attendance, etc.).
- Hidden layers: 2 or 3 layers with ReLU activations.
- Output layer: Performance prediction (probability of belonging to each category).

The application of this model would be to predict multiple academic performance metrics with high accuracy (LeCun et al., 2015).

The advantage of using this method is its high performance in complex tasks with large volumes of data.

Model training and evaluation

The data will be split as follows:

- 70% of the data will be used for training the model.
- 30% of the data will be used for testing.

The model evaluation metrics will include:

- Accuracy: Percentage of correct predictions (Gevorgyan, 2025).
- Recall and precision: Evaluate the balance between false positives and negatives.

DOI https://doi.org/10.22533/at.ed.82082126020112

- F1 Score: Harmonic mean between precision and recall (Powers, 2011).
- Confusion Matrix: Visualize correct and incorrect classifications (Sammut & Webb, 2010).
- Mean absolute error (MAE): For regression models, measures the average magnitude of errors (Willmott & Matsuura, 2005).

It is necessary to perform cross-validation of the data, using K-fold cross-validation (Refaeilzadeh et al., 2009) to ensure the model is robust and generalizes well to new data.

Implementation of the academic performance management algorithm

We should follow these steps for implementation:

- Integration into a Learning Management System (LMS): The algorithm can generate personalized alerts for at-risk students (Zhang et al., 2021).
- Monitoring dashboards: Visualize key metrics such as model predictions, influential variables, and intervention rates.
- Automated interventions: Recommend tutoring, additional materials, or specific resources to students with projected low performance (LeCun et al., 2015).

Practical example of using academic performance management prediction algorithm

A university implements a system to manage the academic performance of its students.

- Input: A student's academic history, tutoring attendance, LMS usage, and demographic data.
- Output: The model predicts that the student has an 85% probability of achieving low performance.
- Action: The university automatically assigns personalized tutoring and sends notifications about academic improvement strategies.

Blockchain: A technology for the traceability and security of educational data in universities

The rise of emerging technologies, such as blockchain and predictive algorithms, offers opportunities to improve educational and administrative processes in universities (Tapscott & Tapscott, 2016). Blockchain technology provides a decentralized, secure, and immutable system for managing data (Jaime, 2019). Some of its most relevant applications in the university context are:

1. Management of academic credentials:
   - Issuance of immutable digital degrees, certificates, and grades (Jaime, 2020).
   - Instant and secure verification of credentials by third parties (employers, institutions).

2. Management of academic records:
   - Decentralized recording of student data, such as completed courses, attendance, participation in extracurricular activities, and achievements (Maestre et al., 2023).

3. Financial transactions:

- Automation of tuition and scholarship payments through smart contracts (Chen et al., 2018).

3. Research traceability:

- Ensuring intellectual property and authenticity of university publications and projects (Yli-Huumo et al., 2016).

4. Management of Digital identities:

- Providing each student with a unique and verifiable digital identity (Gkontzis et al., 2019).

Integration of blockchain technology and predictive AI algorithms in universities

The combination of blockchain and predictive algorithms allows overcoming challenges related to the quality, security, and privacy of educational data. This integration ensures that predictive algorithms work with accurate and reliable information, while respecting privacy and optimizing university processes (Chen et al., 2018).

Management of university student data

Blockchain can act as a secure and decentralized repository for storing student data, such as:

- Grades.
- Attendance records.
- Exam results.
- Use of virtual learning platforms (LMS).

The advantages that blockchain technology brings to predictive algorithms include:

1. Data integrity: Data stored on blockchain is immutable, eliminating the risk of manipulation or corruption (Rani, Sachan & Kukreja, 2023).
2. Single source: Algorithms can access a reliable data record to make accurate predictions (Sharples & Domingue, 2016).
3. Real-Time updates: Changes in students' academic progress can be reflected on blockchain and immediately used by the algorithms (Gkontzis et al., 2019).

One use of blockchain technology combined with predictive algorithms would be to analyze the academic history stored on blockchain to identify patterns suggesting a risk of dropout. The university can quickly intervene with tutoring or support programs (Alammary et al., 2019).

Personalization of learning

Predictive algorithms analyse student data to recommend personalized learning strategies. Blockchain technology facilitates this personalization using digital student identities, which contain relevant information such as:

- Preferred learning methods.
- Resources used on LMS platforms.
- Previous assessments.

The operation of such personalized learning using predictive algorithms combined with blockchain technology would be:

1. Blockchain stores the student's academic data and preferences.

DOI https://doi.org/10.22533/at.ed.8208126020112

2. Predictive algorithms process this data and suggest:

- Personalized courses.
- Additional resources.
- Activities to enhance learning.

The key benefit of combining these two technologies is that personalization improves the educational experience and fosters better academic performance (Rani, Sachan & Kukreja, 2023).

Automation through smart contracts

Smart contracts on blockchain allow automating processes related to predictive interventions, eliminating the need for manual intervention (Chen et al., 2018).

A practical example of combining blockchain technology and predictive algorithms for automation through smart contracts could be as follows:

1. A predictive algorithm detects that a student is at high risk of dropping out of university.
2. A smart contract is automatically triggered to:

- Notify the student and tutors.
- Release resources such as scholarships or personalized counseling.
- Record the intervention on blockchain to ensure traceability.

The impact of combining these two technologies is that automation ensures interventions are quick and effective, optimizing university resources (Sharples & Domingue, 2016).

Ethics, privacy and transparency

One of the biggest challenges in using predictive algorithms is the lack of transparency in the decisions they make. Blockchain technology can address this issue by recording in its system:

- The decisions made by the algorithms.
- The variables used to make predictions.

The benefits that blockchain technology combined with predictive algorithms could bring include:

- Auditability: Students can verify how and why decisions were made based on predictions (Gkontzis et al., 2019).
- Privacy: Blockchain allows students to control what data they share with the algorithms (Alammary et al., 2019).
- Regulatory compliance: The integration ensures that universities comply with regulations such as GDPR, protecting sensitive data (Rani, Sachan & Kukreja, 2023).

Results. Projects in the implementation of blockchain technology and predictive AI algorithms at the university

This section reports the multicase analysis by mapping (i) which UCAS layers are present in each institution, (ii) how multi-source student data are used for prediction and decision-making, and (iii) how blockchain is currently deployed for credential assurance or governance. The analysis is interpreted in light of the established learning-analytics and blockchain-in-education literatures, which emphasize individualized risk prediction and auditable credential infrastructures, respectively (Romero & Ventura, 2020; Alammary et al., 2019; Sharples & Domingue, 2016).

Comparative Case Overview

Table 1 summarizes core UCAS-aligned variables per case: integrated data sources for prediction, targeted outcomes, blockchain deployment focus, degree of automated orchestration, and primary institutional priority.

Two immediate observations arise. First, the most mature multi-source predictive uses (academic, engagement and financial) are found where retention/revenue pressures are explicit (Public Research University; Digital-First University), aligning with evidence that combining heterogeneous signals improves educational risk detection (Zhang et al., 2021; Almalawi et al., 2024). Second, blockchain deployments cluster around verifiable credentials and institutional signaling (e.g., Melbourne's micro-credential pilot), consistent with the educational blockchain literature (University of Melbourne, 2017; Sharples & Domingue, 2016).

Cross case Analysis

Degree of data fusion

A clear gradient appears in data integration:

- High fusion of academic + financial + engagement data. Public Research University and Digital-First University fuse tuition/payment timing, LMS inactivity, missed administrative steps, and performance indicators into individualized alerts that trigger retention and financial-aid actions—a pattern consistent with best-practice early-warning models (Zhang et al., 2021; Romero & Ventura, 2020).
- Academic/engagement integration with limited financial variables. Stanford and MIT emphasize progression modeling and personalization but treat financial/administrative signals in separate systems, a siloing commonly noted in implementations preceding full data governance alignment (Alammary et al., 2019).
- Consumer-behavior orientation. The University of Nicosia tracks program choices and payment behavior as market signals for pathway personalization and conversion—an approach coherent with blockchain-credential ecosystems aimed at student-owned records and portability (Sharples & Domingue, 2016).
- Lifelong-learner horizon. Melbourne focuses on stackable micro-credentials and re-engagement over time, leveraging credential portability to shape long-term learner trajectories (University of Melbourne, 2017).

Predictive intelligence and intervention logic

All institutions deploy predictive/early-warning logic, but for distinct managerial purposes:

- Retention and risk mitigation. Public Research University and Stan-

DOI https://doi.org/10.22533/at.ed.82082126020112

Institution	Integrated Data Sources Used for Prediction	Targeted Predictive / Decision Outcome	Blockchain Deployment Focus	Level of Automated Orchestration / Smart Contract Like Triggers	Primary Institutional Priority
MIT	Academic performance, program progression, engagement with digital learning platforms; limited financial behavior integration	Early identification of academic risk and progression bottlenecks; personalized pathway recommendations	Verifiable digital credentials and tamper-evident records of achievement	Low–moderate: blockchain mainly for credential issuance; intervention workflows largely manual	Academic excellence, signaling of credential trust and portability
Stanford University	Academic metrics, LMS engagement data, advising interactions, support service usage; emerging attention to well-being / support-seeking patterns	Dropout / stop out risk, time to degree optimization, allocation of advising resources	Exploratory work on secure data sharing and auditability rather than large-scale credential issuance	Moderate: alerts routed to advising/success offices; escalation protocols semi-formal	Student success and retention, advisor efficiency
University of Nicosia	Enrollment behavior, program selection decisions, payment timing, interaction with decentralized learning platforms	Enrollment conversion, payment reliability, pathway personalization for specialized blockchain/crypto programs	Full blockchain credentialing; on-chain proof of completion; emphasis on student-owned records	Moderate–high: smart-contract logic for credential issuance and verification is institutionalized	Credential portability, global signaling, revenue through specialization niches
University of Melbourne	Academic progress, micro-credential uptake, continuing education behavior, career services engagement	Lifelong learning pathway modeling, micro-credential stacking, learner re-engagement after graduation	Blockchain-backed / ledger-backed digital certificates for micro-credentials and short courses	Moderate: partial automation of issuing recognized credentials; limited automation of retention interventions	Lifelong learner re-engagement and brand reputation
Public Research University	Academic performance, attendance patterns, tuition payment behavior, missed deadlines, financial aid usage, LMS disengagement signals	Early dropout risk prediction, proactive retention outreach, targeted financial intervention (e.g., revised payment plans)	No production blockchain; governance handled through internal compliance/ audit offices	High for analytics triggers: risk alerts auto-generate advisor outreach tasks and financial aid review	Retention and revenue stability
Digital-First University	Clickstream data from online platforms, pacing in selfdirected modules, payment installment behavior, support ticket history	Continuation probability per module, churn prediction between modules, upsell to next certificate / specialization track	Blockchain primarily for portable verifiable certificates and proof of skills for employability	Moderate–high: intervention workflows partially automated (payment plan offers, outreach nudges)	Scalable personalization and conversion efficiency

Table 1. Cross case comparison of UCAS relevant capabilities

ford route alerts to advising/tutoring/financial-aid with defined service expectations—consistent with literature linking predictive alerts to operational student-success workflows (Zhang et al., 2021).

- Pathway steering and demand shaping. MIT and Digital-First University use prediction to recommend feasible pathways and progression, a use aligned with program-design and resource-allocation perspectives in learning analytics (Romero & Ventura, 2020).
- Product/portfolio strategy. Melbourne and Nicosia analyze which credential formats drive persistence and perceived value, resonating with blockchain-enabled micro-credential ecosystems (University of Melbourne, 2017; Alammary et al., 2019).

Blockchain as Governance and Orchestration Layer

Across cases, blockchain is most mature in credential integrity and portability—e.g., verifiable certificates and micro-credentials—rather than in internal orchestration of predictive interventions. This pattern mirrors the broader field's current emphasis (Sharples & Domingue, 2016; Alammary et al., 2019).

Selective signs of internal auditability appear (timestamped actions, immutable logs), but smart-contract-driven cross-unit triggers remain rare in production. Where automation exists (e.g., outreach/job tickets, payment-plan offers), it is typically CRM-based rather than ledger-governed, despite the potential of smart contracts to encode rules and reduce discretion (Casino et al., 2019; Chen et al., 2018).

Organizational capabilities

Three recurrent enablers distinguish operational from pilot-level practice:

- Cross-unit data governance bodies that align enrollment management, academic affairs, student success, and finance—turning analytics from dashboards into task assignment (Romero & Ventura, 2020).
- Defined escalation protocols for advisor outreach and documentation, which are prerequisites for any move toward verifiable, auditable intervention workflows (Casino et al., 2019).
- Executive framing of analytics as strategy, especially in micro-credential ecosystems that depend on verifiable records and external trust (University of Melbourne, 2017; Sharples & Domingue, 2016).

Key Findings

Finding 1. Partial UCAS implementations are common; full end-to-end integration is not yet observed.

Institutions tend to be strong in one or two layers (e.g., predictive + workflows; or credential blockchain) but do not yet connect predictive outputs to a blockchain-governed intervention layer. This reflects the current state of the field (Sharples & Domingue, 2016; Alammary et al., 2019).

Finding 2. High-performing predictive use cases treat students as both learners and consumers, integrating financial and engagement signals with academics.

DOI https://doi.org/10.22533/at.ed.8208126020112

Where tuition/payment and engagement signals are fused with academic indicators, institutions report earlier and more actionable risk detection, consistent with empirical reviews (Zhang et al., 2021; Almalawi et al., 2024).

Finding 3. Blockchain is currently leveraged primarily for external trust (verifiable credentials), not yet for internal accountability of data-driven interventions.

Credential assurance is production--ready; orchestration via smart contracts remains emergent (University of Melbourne, 2017; Sharples & Domingue, 2016; Casino et al., 2019).

Finding 4. Formalized intervention workflows convert predictive alerts into managerial infrastructure.

Defined service-level expectations (who acts, when, how logged) are the decisive step from pilots to operations, aligning with implementation guidance in the learning-analytics literature (Romero & Ventura, 2020).

Finding 5. Organizational capability is the principal bottleneck to UCAS realization.

AI methods and even blockchain pilots are available, but without cross-unit governance and auditable workflows, predictive insights remain siloed (Chen et al., 2018; Casino et al., 2019).

Discussion. Future research on university consumer behavior combining predictive algorithms and blockchain technology

The combination of predictive algorithms and blockchain not only transforms the current experience but also creates new opportunities for the future:

Competency based education: Blockchain can store micro-credentials obtained in specific courses, allowing students to demonstrate acquired skills in a granular manner (González and Fernández, 2022).

Educational marketplace: With blockchain, universities could create decentralized platforms where students select courses, professors, and specific resources according to their needs, democratizing education (Maestre et al., 2023).

Holistic predictions: As algorithms are fed more data (academic, social, financial), predictions will become more accurate, enabling increasingly effective intervention strategies (Maestre et al., 2023).

Expansion of decentralized educational networks: More universities could join global blockchain networks, creating an interoperable educational ecosystem (Xu, 2024).

Universal portable credentials: Students will be able to easily transfer their academic records between institutions without administrative barriers (Mata & Cruz, 2022).

Micro credentials: Blockchain will allow the issuance of certificates for specific courses or acquired skills, promoting continuous learning (Maestre et al., 2023).

Conclusions

The integration of predictive algorithms and blockchain technology in the university environment not only transforms how institutions manage their internal processes but also redefines the experience of the university consumer, i.e., the students.

These technological tools offer advanced solutions to personalize, predict, and optimize interactions between students and universities. Below are key conclusions derived from the analysis of their applications and benefits, accompanied by relevant references.

Transformation in the educational experience of the University consumer

Personalization of learning

Predictive algorithms allow the identification of patterns in student behavior through the analysis of large volumes of data, such as:

- Participation in classes and digital platforms.
- Grades.
- Learning preferences. This leads to the creation of personalized educational experiences, tailored to the individual needs of each student.

The impact has been that personalization improves student satisfaction and engagement, significantly reducing the university dropout rate.

Evidence according to Romero and Ventura (2020), the use of predictive models in education improves academic performance and increases motivation by providing tailored recommendations.

Improvement in student decisión - making

By offering predictions about student performance and possible academic trajectories, students can make more informed decisions about courses, specializations, and career goals.

The impact has been that improved student decision-making positions universities as proactive agents in their students' success, increasing their confidence in the institution.

Reduction of student dropout through early risk identification

Predictive algorithms can detect signs of potential student dropout, such as:

- Lack of engagement in academic activities.
- Poor academic performance.
- Economic or personal problems reflected in external data.

The impact has been that by implementing intervention strategies based on these models, universities can offer personalized tutoring, financial support, and counseling to at-risk students.

Evidence from Zhang et al. (2020) observed a 25% reduction in dropout rates at universities that employed predictive analytics to develop support programs.

Trust and transparency through blockchain technology

Secure and transparent dredential management

Blockchain technology allows the recording of academic credentials such as degrees, certificates, and student achievements on an immutable network. This eliminates forgery and facilitates verification by employers. The impact has been that students have greater control over their data, which builds trust in institutions. Evidence from Sharples and Domingue (2016) highlights that blockchain significantly improves trust in academic management systems.

DOI https://doi.org/10.22533/at.ed.8208126020112

Protection of personal data

Blockchain technology, being decentralized, ensures that students can decide who has access to their data and under what conditions, complying with regulations such as GDPR. The impact has been the empowerment of the university consumer and strengthens the trust relationship between students and universities.

Operational efficiency and automation

Smart Contracts

Smart contracts automate administrative processes such as:

- Enrollments.
- Scholarship assignments.
- Issuance of certificates.

The impact has been that operational efficiency increases, allowing universities to allocate more resources to learning and innovation.

Resource optimization

By combining blockchain technology with predictive algorithms, universities can foresee future demands, such as tutor assignments or course planning. The impact has been cost reduction and improved educational experience. Evidence from Casino et al. (2019) concludes that blockchain-based automation is key to reducing bureaucracy in the education sector.

References

Alammary, A., Alhazmi, S., Almasri, M., & Gillani, S. (2019). Blockchain-Based Applications in Education: A Systematic Review. *Applied Sciences*, *9*(12), 2400. https://doi.org/10.3390/app9122400

Alexander, G. D. (January 29, 2025). The future of higher ed: Why colleges must invest in AI, the metaverse, and blockchain. *eCampus News*. https://www.ecampusnews.com/digital-innovation/2025/01/29/colleges-must-invest-ai-metaverse-blockchain/

Almalawi, A., Soh, B., Li, A., & Samra, H. (2024). Predictive Models for Educational Purposes: A Systematic Review. *Big Data and Cognitive Computing*, *8*(12), 187. https://doi.org/10.3390/bdcc8120187

Alsaadi, A. & Bamasoud, D.M. (2021). Blockchain technology in Education system. International Journal of Advanced Computer Science and Applications 12(5). https://doi.org/10.14569/IJACSA.2021.0120585

Baker, R. S., & Inventado, P. S. (2014). Educational data mining and learning analytics. In *Learning Analytics* (pp. 61-75). Springer. https://link.springer.com/chapter/10.1007/978-1-4614-3305-7_4

Breiman, L. (2001). Random forests. *Machine Learning*, 45(1), 5-32. https://link.springer.com/article/10.1023/A:1010933404324

Brownlee, J. (2021). *Machine Learning Mastery with Python: Understand Your Data, Create Accurate Models, and Work Projects End-to-End*. Machine Learning Mastery.

Casino, F., Dasaklis, T. K., & Patsakis, C. (2019). "A Systematic Literature Review of Blockchain-Based Applications: Current Status, Classification and Open Issues." *Telematics and Informatics, 36*, 55-81. https://www.sciencedirect.com/science/article/pii/S0736585318306324

Chen, G., Xu, B., Lu, M., & Chen, N. S. (2018). Exploring blockchain technology and its potential applications for education. *Smart Learning Environments*, 5(1), 1-10. https://doi.org/10.1186/s40561-017-0050-x

Cheng J-C, Lee N-Y, Chi C, Chen Y-H (2018) Blockchain and smart contract for digital certificate. *IEEE International Conference on Applied System Invention (ICASI)*, Chiba, Japan, 2018, pp. 1046-1051, doi: 10.1109/ICASI.2018.8394455. https://doi.org/10.1109/ICASI.2018.8394455

Chollet, F. (2018). *Deep Learning with Python*. Manning Publications.

Cutler, D. R., Edwards, T. C., Beard, K. H., & others (2007). *Random forests for classification in ecology*. Ecology, 88(11), 2783–2792. https://esajournals.onlinelibrary.wiley.com/doi/abs/10.1890/07-0539.1

Fernández-Delgado, M., Cernadas, E., Barro, S., & Amorim, D. (2014). Do we need hundreds of classifiers to solve real world classification problems? *Journal of Machine Learning Research*, 15(1), 3133-3181. https://jmlr.org/papers/volume15/delgado14a/delgado14a.pdf

Gevorgyan, A. (2025). Predicting Students' Academic Performance Using Correlation and Regression Analysis. https://doi.org/10.1007/978-3-031-96380-3_25

Gkontzis, A., Kotsiantis, S., Panagiotakopoulos, C. & Verykios, V. (2019). A predictive Analytics Frameworks as countermeasure for attrition of students. Interactive Learning environments. 30. https://doi.org/10.1080/10494820.2019.1674884

Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep Learning*. MIT Press.

Hosmer, D. W., Lemeshow, S., & Sturdivant, R. X. (2013). *Applied Logistic Regression*. Wiley.

Jaime Maestre, R. (mayo de 2019). La Blockchain revolucionarà la comunicació de l'empresa. Comunicació: Revista de Recerca i d'Anàlisis, 36(1), pp 95-116. https://revistes.iec.cat/index.php/TC/issue/view/9802

Jaime Maestre, R. (2020). Un ejemplo de educación financiada mediante criptomoneda: la ICO de la IEBS Business School. Tecnología, Ciencia y Educación, 15, 143-163. https://www.tecnologia-ciencia-educacion.com/index.php/TCE/article/view/375

Kazem, S., Noroozi, O., Van Gingkel, S., Macfadyen, L.P. & Biemans, H. (2022). A systematic review of the role of learning analytics in enhancing feedback practices in higher Education. Educational Research Review vol. 37. https://doi.org/10.1016/j.edurev.2022.100489

Kurni, M., Mohammed, M. & Srinivasa, K. (2023). Predictive Analytics in Education. In Book: A beginner's guide to introduce artificial intelligence in teaching and learning (pp. 55-81). https://link.springer.com/chapter/10.1007/978-3-031-32653-0_4

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, 521(7553), 436-444. https://doi.org/10.1038/nature14539

Liaw, A., & Wiener, M. (2002). Classification and regression by randomForest. *R news*, 2(3), 18-22. https://journal.r-project.org/articles/RN-2002-022/RN-2002-022.pdf

Maestre, R.J., Bermejo Higuera, J., Gámez Gómez, N. *et al.* The application of blockchain algorithms to the management of education certificates. *Evol. Intel.* 16, 1967–1984 (2023). https://doi.org/10.1007/s12065-022-00812-0

Mata, J.M. & Cruz, S. (2022). Blockchain en la educación: su uso en credenciales académicas. Revista Digital Universitaria. 23. https://doi.org/10.22201/cuaieed.16076079e.2022.23.1.1

MIT-IBM Watson AI Lab. (s.f.). About the lab. https://mitibmwatsonailab.mit.edu/

Munir, J., Faiza, M. & Daud, S. (2023). The impact of socio-economic status on academic achievement. Journal of social sciences review. 3. 695-705. https://doi.org/10.54183/jssr.v3i2.308

Nakamoto, S. (2009). Bitcoin: A Peer-to-Peer Electronic Cash System. https://www.researchgate.net/publication/228640975_Bitcoin_A_Peer-to-Peer_Electronic_Cash_System

Nair, V., & Hinton, G. E. (2010). Rectified linear units improve restricted Boltzmann machines. In *Proceedings of the 27th international conference on machine learning (ICML)* (pp. 807-814). https://dl.acm.org/doi/10.5555/3104322.3104425

Ngulube, P., & Ncube, M. M. (2025). Leveraging Learning Analytics to Improve the User Experience of Learning Management Systems in Higher Education Institutions. *Information*, *16*(5), 419. https://doi.org/10.3390/info16050419

Poudel, U. & Gajjela, S. (2019) A blockchain based micro-credentialing sytem. https://doi.org/10.13140/RG.2.2.18403.20008

Powers, D. M. (2011). Evaluation: From precision, recall and F-measure to ROC, informedness, markedness and correlation. *Journal of Machine Learning Technologies*, 2(1), 37-63. https://arxiv.org/abs/2010.16061

Psimopoulos, C. (15 october, 2024). Artificial Intelligence: Bioethical considerations and limitations. *Public Orthodoxy*. https://publicorthodoxy.org/2024/10/15/artificial-intelligence-bioethics/

Rani, P., Sachan, R.K. & Kukreja, S. A (2023) A systematic study on blockchain technology in education: initiatives, products, applications, benefits, challenges and research direction. *Computing* 106, 405–447. https://doi.org/10.1007/s00607-023-01228-z

Refaeilzadeh, P., Tang, L., & Liu, H. (2009). Cross-validation. In *Encyclopedia of Database Systems* (pp. 532-538). Springer. https://link.springer.com/rwe/10.1007/978-0-387-39940-9_565

Romero, C., & Ventura, S. (2020). *Educational data mining: A review of the state of the art*. IEEE Transactions on Systems, Man, and Cybernetics, 40(6), 601-618. https://doi.org/10.1109/TSMCC.2010.2053532

Sammut, C., & Webb, G. I. (2010). *Encyclopedia of Machine Learning*. Springer Science & Business Media. https://doi.org/10.1007/978-0-387-30164-8

Schmidhuber, J. (2014). Deep learning in neural networks: An overview. *Neural Networks*, 61, 85-117. https://doi.org/10.1016/j.neunet.2014.09.003

Sharples, M., & Domingue, J. (2016). The blockchain and kudos: A distributed system for educational record, reputation and reward. *European Conference on Technology Enhanced Learning*, 490-496. https://doi.org/10.1007/978-3-319-45153-4_48

Stanford Artificial Intelligence Laboratory. (s.f.). About the lab. https://ai.stanford.edu/about/

Stanford Center for Blockchain Research. (s.f.). Mission. https://cbr.stanford.edu/

Syed Mustapha, S. M. F. D. (2023). Predictive Analysis of Students' Learning Performance Using Data Mining Techniques: A Comparative Study of Feature Selection Methods. *Applied System Innovation*, *6*(5), 86. https://doi.org/10.3390/asi6050086

Tapscott, D., & Tapscott, A. (2016). *Blockchain Revolution: How the Technology Behind Bitcoin is Changing Money, Business, and the World*. Penguin.

University of Melbourne. (27 April, 2017). Melbourne University to pilot a distributed database for micro-credentials. https://www.unimelb.edu.au/newsroom/news/2017/april/melbourne-university-to-pilot-a-distributed-database-for-micro-credentials

University of Melbourne. (s.f.). AI research initiatives. https://cis.unimelb.edu.au/research/artificial-intelligence/research

Willmott, C. J., & Matsuura, K. (2005). Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance. *Climate Research*, 30(1), 79-82. https://www.jstor.org/stable/24869236

Xu, Zhiyi. (2024). AI in education: Enhancing learning experiences and student outcomes. Applied and Computational Engineering. 51. 104-111. https://doi.org/10.54254/2755-2721/51/20241187

Yli-Huumo, J., Ko, D., Choi, S., Park, S., & Smolander, K. (2016). Where is current research on blockchain technology? - A systematic review. *PLOS ONE*, 11(10), https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0163477

Zhang, Yupei & Yun, Yue & An, Rui & Cui, Jiaqi & Dai, Huan & Shang, Xuequn. (2021). Educational Data Mining Techniques for Student Performance Prediction: Method Review and Comparison Analysis. Frontiers in Psychology. 12. https://doi.org/10.3389/fpsyg.2021.698490

Zheng, Z., Xie, S., Dai, H., Chen, X., & Wang, H. (2018). Blockchain challenges and opportunities: A survey. *International Journal of Web and Grid Services*, 14(4), 352-375. https://www.henrylab.net/wp-content/uploads/2017/10/blockchain.pdf