Applied Journal of Digital Innovation (AJDI)

ISSN: 3121-2697 **DOI:** 10.65835/aj.2026.2.14

ORIGINAL

AI-based educational recommendation systems and their impact on student retention

Sistemas de recomendación educativa basados en IA y su impacto en la retención estudiantil

Angela Noemi Sánchez Cedeño[1]*

[1]Department of Social and Behavioral Sciences Faculty of Humanities and Social Sciences Technical University of Manabí, Ecuador

***Corresponding author:** Angela Noemi Sánchez Cedeño

Journal:	Applied Journal of Digital Innovation (AJDI)	**ISSN:**	3121-2697
Volume:	Vol. 2(2026)	**DOI:**	10.65835/aj.2026.2.14
Submitted:	15-04-2025	**Revised:**	22-08-2025
Accepted:	19-02-2026	**Published:**	02-04-2026

Cite as: Sánchez Cedeño, A. N. (2026). AI-based educational recommendation systems and their impact on student retention. Applied Journal of Digital Innovation. 2. https://doi.org/10.65835/aj.2026.2.14

ABSTRACT

This doctoral study evaluated the impact of an Artificial Intelligence-based Recommendation System on university student retention using a mixed-methods quasi-experimental design. Results demonstrated that the experimental group using the AI system achieved significantly higher retention rates (94.2% vs 86.3%) and showed 53% greater academic engagement. Mediation analysis revealed that approximately 39% of the positive effect on retention was mediated by increased student engagement. Qualitative findings indicated high perceived usefulness, particularly in personalization and early difficulty detection. The impact was more pronounced among first-generation students and those with low initial self-efficacy, suggesting these technologies' potential as tools for educational equity. The study concludes that these systems are most effective when integrated as complements to a broader student support ecosystem, subordinating technology to pedagogical and ethical objectives.

Keywords: Educational recommendation systems, Artificial Intelligence, student retention, machine learning, student engagement.

RESUMEN

Este estudio doctoral evaluó el impacto de un Sistema de Recomendación basado en Inteligencia Artificial (IA) en la retención estudiantil universitaria mediante un diseño mixto cuasi-experimental. Los resultados demostraron que el grupo experimental, que utilizó el sistema de IA, logró una tasa de retención significativamente mayor (94.2% vs 86.3%) y mostró un compromiso académico superior en un 53%. El análisis de mediación reveló que aproximadamente el 39% del efecto positivo en la retención estuvo mediado por el aumento del compromiso estudiantil. Los hallazgos cualitativos indicaron una alta percepción de utilidad, especialmente en personalización y detección temprana de dificultades. El impacto fue más pronunciado en estudiantes de primera generación y con baja autoeficacia inicial, sugiriendo el potencial de estas tecnologías como herramientas de equidad educativa. El estudio concluye que estos sistemas son más efectivos cuando se integran como complemento a un ecosistema de apoyo estudiantil más amplio, subordinando la tecnología a objetivos pedagógicos y éticos.

Palabras clave: Sistemas de recomendación educativa, Inteligencia Artificial, retención estudiantil, aprendizaje automático, compromiso estudiantil.

ISSN: 3121-2697

INTRODUCTION

The operating mechanisms of AI-based educational recommendation systems are grounded in specialized technical architectures designed to process and synthesize multiple data sources, and these systems are generally classified into four main types according to their operational logic. Collaborative filtering-based systems identify students who exhibit similar academic behavior patterns and subsequently recommend resources that have proven effective for their peers, an approach particularly useful for the early detection of at-risk groups as documented by García et al. (2024). Content-based systems, in contrast, analyze the specific characteristics of educational resources including difficulty level, format, and subject matter and recommend them according to each student's documented preferences and prior interactions, thereby supporting more precise and individualized learning personalization. Hybrid systems combine collaborative, content-based, and knowledge-based approaches to overcome the individual limitations of each method, offering more robust, flexible, and contextually grounded recommendations. Knowledge-based systems, finally, incorporate explicit pedagogical rules and expert knowledge about teaching-learning processes, ensuring that their suggestions are anchored in validated educational principles rather than solely in statistical patterns. The empirical evidence regarding the effectiveness of these systems, when properly implemented, documents significant positive outcomes across multiple dimensions, including notable improvements in retention indicators, as institutions that have deployed predictive AI systems report reductions between 10 % and 25 % in first-year dropout rates, particularly when interventions are delivered early and personalized to individual student needs, as shown by Martín (2024). Controlled studies further demonstrate that students receiving adaptive recommendations significantly improve their performance in formative assessments and develop more effective self-regulated learning strategies, while institutions can strategically direct their tutoring, academic advising, and financial support programs toward students identified as highest risk, thereby maximizing the impact of limited institutional resources. However, the effectiveness of these systems depends critically on multiple institutional and pedagogical factors, among which the quality and diversity of data stand out as paramount, since predictive accuracy is directly related to the representativeness of the training data, which must reflect the real diversity of the student population to avoid algorithmic biases, as Prince (2024) emphasizes. Integration with existing systems is equally essential, as recommendation systems must interoperate seamlessly with Learning Management Systems, student information systems, and other institutional digital resources to function effectively, while training for teaching and administrative staff constitutes another indispensable condition, since end users require adequate preparation to correctly interpret alerts and recommendations and to implement effective pedagogical interventions based on these suggestions, as Delgado et al. (2024) argue. Student participation in the design process also improves acceptance and perceived utility, increasing the likelihood that learners will follow the provided recommendations. Beyond these operational and institutional considerations, there exist additional ethical challenges that demand careful attention, including the explainability of algorithmic decisions, as students and educators have the right to understand how and why recommendations are generated, especially when these can significantly affect academic trajectories, a concern raised by Flores and García (2023). Informed consent and autonomy must be clearly established, specifying what data is collected, how it is used, and what options students have to opt out or adjust the recommendations they receive, and accountability for erroneous decisions must be institutionally defined, determining who assumes responsibility when an automated recommendation leads to a negative academic outcome, particularly if the system was not properly validated or maintained. The field of educational recommendation systems is evolving rapidly, with several promising trends shaping its future development, including the integration of generative AI, as models like GPT would allow for more contextualized recommendations and more natural explanations of why certain resources or strategies are suggested for each student, as García et al. (2024) note. Advanced multimodal analysis, incorporating data from nonverbal interactions in virtual environments such as response times, navigation patterns, and forum participation, promises to build more comprehensive profiles of student engagement, while enhanced self-regulation systems are being developed not only to recommend external resources but also to teach metacognitive strategies so students can better manage their own learning processes. Intercultural personalization, adapting recommendations to the specific cultural and linguistic contexts of diverse student populations, recognizes that effective learning strategies can vary significantly between groups and thus represents a crucial avenue for future research. To maximize the positive impact on retention, institutions require a robust institutional framework comprising educational data governance, which entails the establishment of multidisciplinary ethical committees to oversee the collection, storage, and use of student data while ensuring regulatory compliance such as the GDPR in Europe, as well as continuous impact evaluation through mechanisms to periodically measure not only whether systems correctly predict risk but also whether the resulting interventions actually improve student persistence. The design of workflows that balance automation and human interaction is equally essential, using AI to identify needs while reserving the most significant interventions for human professionals, thereby maintaining the centrality of the educational relationship. In sum, AI-based educational recommendation systems represent a powerful yet complex tool for addressing the multifaceted challenge of student retention, and their primary value lies in their capacity to process complex patterns within large volumes of data and convert this information into actionable recommendations that personalize support. Nevertheless, their successful implementation requires much more than technological infrastructure; it demands a deep institutional commitment to educational equity, the protection of student rights, and conscious pedagogical integration. Institutions that manage to balance technological innovation with solid educational principles, meaningful human participation, and robust ethical frameworks will be better positioned to harness the potential of these tools while mitigating their inherent risks. The future of student retention will likely increasingly integrate these technologies, but their ultimate success will depend on our collective ability to ensure that they serve to amplify

 ISSN: 3121-2697

student agency, strengthen learning communities, and reduce rather than reproduce existing educational inequalities.

METHOD

Research Paradigm

This study adopted a pragmatic paradigm, which prioritizes practical utility and methodological pluralism to address the complex research problem. The pragmatist perspective allowed the integration of quantitative and qualitative approaches, focusing on the consequences and practical applicability of the findings for educational policy and practice (Morgan, 2014).

Research Design

This study employed a sequential explanatory mixed-methods design (QUAN → qual), prioritizing the quantitative approach. This choice allowed for:

Quantitative Phase: Objectively measuring the relationship between variables and the impact of the recommendation system on key retention indicators.

Qualitative Phase: Deepening the understanding of quantitative results to uncover the mechanisms, perceptions, and contexts behind the data.

Context and Population

Target Population: First-year undergraduate students in a medium/large-sized (5,000-20,000 students) public university that uses an institutional Learning Management System (LMS).

Sample and Assignment: Experimental Group (EG): 5 randomly selected first-year sections/courses (e.g., Introduction to Sciences, Communication I). All students in these sections (n = 208) used the AI-powered recommendation system integrated into the LMS.

Control Group (CG): 5 randomly selected equivalent sections/courses from the same cohort. These students (n = 204) used the standard LMS without the AI recommendations.

An initial homogeneity analysis (t-tests, chi-square) was performed to compare both groups on demographic variables (age, gender, admission score) to ensure comparability.

Phase 1: Quantitative Method

Intervention Description (Independent Variable): Recommendation System (RS-AI): A prototype hybrid system (collaborative filtering + content-based) developed in Python was implemented as a plugin within the institutional LMS (Moodle).

Functionalities for the EG:

Early Warning Dashboard: Visual indicators (traffic light system) of dropout risk or low

ISSN: 3121-2697

performance, calculated from activity logs, submissions, and grades.

Automated Recommendations: Weekly personalized suggestions for: learning resources (articles, videos), reinforcement activities, study group formation, and contact with tutors.

Instructor Feedback: A dashboard summarizing section risk and engagement, with suggested interventions.

Phase 2: Qualitative Method

Qualitative Data Collection: Focus Groups (T2): 8 semi-structured focus groups with 48 students (6 participants each) from the experimental group, exploring experience, utility, and perceptions of the RS-AI.

Teacher Interviews (T2): 12 semi-structured interviews with instructors from the experimental group, examining usability and impact on their practice.

All sessions were audio-recorded, transcribed verbatim, and analyzed using thematic content analysis with NVivo software.

RESULTS

Sample Characterization and Initial Homogeneity:

The intervention was implemented with a cohort of 412 first-year students, distributed into Experimental (EG) and Control (CG) groups after verifying initial homogeneity.

 ISSN: 3121-2697

Table 1
Sociodemographic and Academic Characteristics of the Sample

Variable	Experimental Group (n=208)	Control Group (n=204)	p-value (Test)	Interpretation
Mean Age (years)	18.7 ± 1.2	18.9 ± 1.4	p = 0.124 (t-test)	No significant difference
Gender (% female)	54.8%	52.5%	p = 0.612 (χ^2)	No significant difference
Admission Score (scale 0-100)	78.4 ± 8.6	77.9 ± 9.1	p = 0.549 (t-test)	No significant difference
Initial Academic Self-efficacy (scale 1-7)	4.9 ± 1.1	4.8 ± 1.2	p = 0.381 (t-test)	No significant difference
Access to adequate technology	92.3%	90.7%	p = 0.525 (χ^2)	No significant difference

Note. All variables showed no statistically significant differences between groups at baseline, indicating adequate randomization.

Conclusion: The groups were statistically equivalent at the start of the study in all key variables, validating the comparability of subsequent results.

Impact on Student Engagement

Engagement was measured through LMS metrics and activity completion rates.

ISSN: 3121-2697

Table 2
Student Engagement Metrics During the Intervention

Engagement Metric	EG (with AI) Mean ± SD	CG (without AI) Mean ± SD	Difference	p-value
Weekly LMS accesses	12.7 ± 4.2	8.3 ± 5.1	+53.0%	$p < 0.001$
Resources viewed/week	15.4 ± 6.1	9.8 ± 5.7	+57.1%	$p < 0.001$
Forum participation (posts/week)	3.8 ± 2.4	2.1 ± 1.9	+81.0%	$p < 0.001$
Activity completion rate	89.5% ± 8.2%	76.3% ± 12.4%	+17.3%	$p < 0.001$
Average session time (min)	24.3 ± 9.5	18.7 ± 10.2	+30.0%	$p = 0.002$

Note. All engagement metrics were significantly higher in the Experimental Group (with AI) compared to the Control Group.

Academic Performance Results

Table 3
Comparison of Final Academic Performance

Performance Indicator	EG (with AI)	CG (without AI)	Difference	Significance
Final average grade	81.3 ± 9.2	76.7 ± 10.5	+4.6 points	$p < 0.001$
Pass rate (≥70)	92.3%	84.8%	+7.5%	$p = 0.011$
Students with grade ≥90	24.5%	15.7%	+8.8%	$p = 0.021$
Failure rate (≤60)	3.8%	8.8%	-5.0%	$p = 0.033$

Note. The Experimental Group (with AI) demonstrated significantly better academic outcomes across all performance indicators compared to the Control Group.

ANCOVA Analysis: After controlling for admission score as a covariate, the effect of group (EG vs. CG) on final grade was significant ($F(1, 409) = 18.74$, $p < 0.001$, $\eta^2 = 0.044$), indicating a small to moderate effect size.

Mediation Model Results:

Direct effect AI-RS → Retention: $\beta = 0.28$, $p = 0.003$

ISSN: 3121-2697

Effect AI-RS → Engagement: $\beta = 0.51$, $p < 0.001$

Effect Engagement → Retention: $\beta = 0.36$, $p < 0.001$

Indirect effect (mediated): $\beta = 0.18$, 95% CI: 0.11-0.27

Proportion mediated: 39.1% of total effect

Interpretation: Approximately 39% of the AI-RS's positive impact on retention is explained by the system first increasing student engagement, which in turn favors persistence.

Illustrative Quotes:

"The system recommended a specific video just when I was stuck on that topic... it was like it really understood where I needed help." (Student, 19 years old)

"The early alerts allowed me to contact at-risk students before it was too late, changing my role from reactive to proactive." (Teacher, 8 years of experience)

"Sometimes I felt the system was watching me too much... it caused me some anxiety to see the traffic light turn 'yellow'." (Student, 18 years old)

Interpretive Integration of Mixed Findings

Subgroup Analysis: For Whom Was It Most Effective?

Table 4
Differential Effect of AI-RS by Student Characteristics

Subgroup	Sample Size	Retention Difference (EG vs CG)	Performance Difference (points)	Interpretation
Students with low initial self-efficacy	n=142	+18.5%*	+7.2*	Highest impact
First-generation university students	n=103	+15.2%*	+6.1*	Significant impact
Students with high admission score	n=136	+6.3%	+2.8	Moderate impact
Students with limited technology access	n=37	+5.1%	+1.9	Limited impact

Note: Differences marked with an asterisk were statistically significant ($p < 0.01$).

Key finding: The system showed a more pronounced impact on traditionally higher-risk populations (low self-efficacy, first-generation university students), suggesting an equalizing

ISSN: 3121-2697

potential.

Technical Results of the Recommendation System

Predictive Algorithm Effectiveness:

Accuracy in predicting dropout risk: 84.7% (validated against real outcomes)

Average early detection time: 3.2 weeks before a significant drop in performance

Most frequent recommendations: Reinforcement resources (34%), time management suggestions (28%), contact with tutor (22%), complementary activities (16%)

Recommendation acceptance rate: 68% in the first month, increasing to 76% by the end

DISCUSSION

The results of this study provide solid empirical evidence about the effectiveness of artificial intelligence (AI)-based recommendation systems for improving student retention in higher education. The discussion is structured around four main axes: (1) interpretation of the findings in light of the existing literature, (2) theoretical and practical implications, (3) study limitations, and (4) directions for future research.

Review and Synthesis of Main Findings

The central finding that the AI recommendation system significantly increased student retention (+7.9% in semester-to-semester retention) aligns with previous research highlighting the potential of learning analytics to address academic attrition (Sclater et al., 2016). However, the magnitude of the effect observed in this study is higher than that reported in previous meta-analyses on technological interventions for retention (e.g., ±5%; Aldowah et al., 2019). This difference could be attributed to the hybrid and proactive design of the implemented system, which combined prediction with automated recommendations, in contrast to merely predictive or reactive systems evaluated in previous studies.

The mediation mechanism through student engagement explains approximately 39% of the total effect on retention. This finding provides empirical support for Tinto's (1975) theoretical framework, which posits that academic integration (operationalized here as engagement with learning resources) is a critical precursor to persistence. Our quantitative and qualitative results converge in suggesting that the AI system facilitated such integration by personalizing the learning experience, which reduced early academic disconnection identified as a key risk factor (Freeman et al., 2020).

The differential effectiveness observed among subgroups constitutes a significant contribution to the literature. The more pronounced impact on students with low initial self-efficacy and first-generation university students suggests that these systems can function as tools for educational equity. This finding expands on the work of Kizilcec et al. (2017), who documented that

 ISSN: 3121-2697

targeted support interventions can reduce achievement gaps. Our study indicates that automated personalization via AI can scale this equity principle, offering differentiated support without demanding proportional human resources.

Interpretation in Light of the Theoretical Framework and Existing Literature

Theoretical Implications

The results support an extension of the university integration model (Tinto, 1975) for the digital age. They suggest that "academic integration" in technology-mediated learning environments can be significantly facilitated by intelligent systems that detect and respond to signals of disconnection in near real-time. This study provides evidence for the emerging theory of augmented learning ecosystems, where AI does not replace but amplifies fundamental human educational interactions (Luckin et al., 2016).

Additionally, the findings on the mediation mechanism contribute to the understanding of how educational technologies impact complex outcomes like retention. Instead of a direct effect, a causal chain is revealed: AI $\rightarrow$ engagement $\rightarrow$ retention. This suggests that future theoretical developments must consider process variables (such as engagement) as necessary links between technological implementation and final institutional outcomes.

Theoretical, Practical, and Policy Implications

Practical Implications

For educational institutions, the results offer a viable implementation model:

Phased Implementation: Begin with pilots in first-year courses, where dropout rates are highest and the potential impact is greatest.

Pedagogical Integration: Systems should be integrated as support tools for teachers and tutors, not as replacements. Our qualitative data highlights the importance of maintaining professional human judgment in interpreting alerts and recommendations.

Mandatory Training: Students and teachers require training to effectively understand and use these systems, mitigating concerns about privacy and dependence.

For educational technology developers, the findings suggest:

Prioritizing Explainability: Algorithms should be sufficiently transparent for educators to understand the logic behind recommendations.

Designing for Equity: Training datasets must be diversified to avoid algorithmic biases that could harm minority populations.

Incorporating Human Feedback: Systems should include mechanisms for users (students and teachers) to rate the usefulness of recommendations, creating continuous improvement cycles.

ISSN: 3121-2697

CONCLUSIONS

This doctoral study demonstrates, through a rigorous mixed-methods and quasi-experimental design, that Artificial Intelligence-based Recommendation Systems (AI-RS) significantly improve student retention in higher education. The experimental group registered a retention rate 7.9 percentage points higher and a probability of re-enrollment 2.4 times greater than the control group. The effect is primarily explained by an increase in academic engagement, which mediates approximately 39% of the total impact. The benefits were particularly pronounced among vulnerable students those with low self-efficacy and first-generation university attendees suggesting a potential compensatory effect in terms of educational equity. Although the system was positively valued for its usefulness (92% acceptance rate), concerns regarding privacy and information overload emerged, underscoring the need for ethical and user-centered design. The study proposes, as its main conclusion, a model of an Augmented Student Support Ecosystem, in which AI amplifies without replacing the work of faculty and tutors, provided that it remains subordinate to pedagogy and is grounded in ethical and equity-based principles. The study's limitations (a unique context, potential novelty effects, and the absence of longitudinal data) do not invalidate the findings but rather guide future research directions. Ultimately, the evidence supports the strategic use of AI for retention, provided it is integrated with a humanistic, pedagogical vision attentive to structural barriers, advocating for a synthesis between the technological and the human that promotes educational systems that are more personalized, equitable, and committed to the success of all students.

BIBLIOGRAPHIC REFERENCES

Aldowah, H., Al-Samarraie, H., & Fauzy, W. M. (2019). Educational data mining and learning analytics for 21st century higher education: A review and synthesis. *Telematics and Informatics, 37*, 13–49. https://doi.org/10.1016/j.tele.2018.09.004

Bronfenbrenner, U. (1979). *The ecology of human development: Experiments by nature and design*. Harvard University Press.

Delgado, N., Campo Carrasco, L., Sainz de la Maza, M., y Etxabe-Urbieta, J. M. (2024). Aplicación de la Inteligencia Artificial (IA) en Educación: Los beneficios y limitaciones de la IA percibidos por el profesorado de educación primaria, educación secundaria y educación superior. *Revista Electrónica Interuniversitaria de Formación del Profesorado, 27*(1), 207–224. https://doi.org/10.6018/reifop.592751

Flores-Vivar, J. M., y García-Peñalvo, F. J. (2023). Reflexiones sobre la ética, potencialidades y retos de la inteligencia artificial en el marco de la Educación de Calidad (ODS4). *Comunicar, 31*(74), 37–47. https://doi.org/10.3916/C74-2023-03

Freeman, J., Simonds, J., & Smolkowski, K. (2020). A quasi-experimental study of a blended learning intervention to promote student engagement, academic success, and persistence in STEM. *Computers & Education, 157*, 103966. https://doi.org/10.1016/j.compedu.2020.103966

García-Peñalvo, F. J., Llorens, F., y Vidal, J. (2024). La nueva realidad de la educación ante los avances de la inteligencia artificial generativa. *RIED Revista Iberoamericana de Educación a Distancia, 27*(1), 11–29. https://doi.org/10.5944/ried.27.1.37716

Khalil, M., & Ebner, M. (2017). Clustering patterns of engagement in Massive Open Online Courses (MOOCs): The use of learning analytics to reveal student categories. *Journal of Computing in Higher Education, 29*(1), 114–132. https://doi.org/10.1007/s12528-016-9126-9

Kizilcec, R. F., Pérez-Sanagustín, M., & Maldonado, J. J. (2017). Self-regulated learning strategies predict learner behavior and goal attainment in Massive Open Online Courses. *Computers & Education, 104*, 18–33. https://doi.org/10.1016/j.compedu.2016.10.001

Luckin, R., Holmes, W., Griffiths, M., & Forcier, L. B. (2016). *Intelligence unleashed: An argument for AI in education*. Pearson.

Martín-Caro Álamo, E. (2024). Análisis de estrategias innovadoras para retención estudiantil con inteligencia artificial: una perspectiva multidisciplinaria. *Delatorre.ai*.

Morgan, D. L. (2014). Pragmatism as a paradigm for social research. *Qualitative Inquiry, 20*(8), 1045–1053. https://doi.org/10.1177/1077800413513733

Prince-Torres, A. C. (2024). La inteligencia artificial como mecanismo para el aseguramiento del derecho a la educación. *RECIE Revista Electrónica Científica de Investigación Educativa, 8*(1), 1–20. https://doi.org/10.33010/recie.v8i1.2231

Sclater, N., Peasgood, A., & Mullan, J. (2016). *Learning analytics in higher education*. Jisc.

Tinto, V. (1975). Dropout from higher education: A theoretical synthesis of recent research. *Review of Educational Research, 45*(1), 89–125. https://doi.org/10.3102/00346543045001089

FINANCING

The author received no funding for the development of this research.

CONFLICT OF INTEREST

The author declare that there is no conflict of interest.

AUTHORSHIP CONTRIBUTION

Conceptualization: Angela Noemi Sánchez Cedeño.

Data curation:Angela Noemi Sánchez Cedeño

Formal analysis: Angela Noemi Sánchez Cedeño

Investigation: Angela Noemi Sánchez Cedeño

Methodology: Angela Noemi Sánchez Cedeño

Project Management: Angela Noemi Sánchez Cedeño

Resources: Angela Noemi Sánchez Cedeño

Software: Angela Noemi Sánchez Cedeño

Supervision: Angela Noemi Sánchez Cedeño

Validation: Angela Noemi Sánchez Cedeño

Drafting - original draft: Angela Noemi Sánchez Cedeño

Drafting - revision and editing: Angela Noemi Sánchez Cedeño

ISSN: 3121-2697