



Revista ISTE Scientist

Vol. 5 Núm. 1 (2026). pp. 16-22

Instituto Tecnológico Superior España

ISSN: 2953-6618

Tipo: Artículo Científico



Como citar: Ortiz-Panezo, J. J. & Sinchiguano-Chiriboga, C. A. (2026).

Aplicación de Modelos de Inteligencia Artificial para la Predicción del Rendimiento Académico en Estudiantes Universitarios de Tecnología de la Información: Application of artificial intelligence models for predicting academic performance in university students of information technology. *ISTE SCIENTIST*, 5(1), 16-22.

<https://revistas.iste.edu.ec/index.php/reviste/article/view/64>

Aplicación de modelos de inteligencia artificial para la predicción del rendimiento académico en estudiantes universitarios de tecnología de la información

Jhon Jairon, Ortiz-Panezo

opjj.1316408200@gmail.com

<https://orcid.org/0009-0004-7089-9573>

Universidad Laica "Eloy Alfaro de Manabí",

Extensión El Carmen

El Carmen - Ecuador

Cesar Augusto, Sinchiguano-Chiriboga

cesar.sinchiguano@uleam.edu.com

<https://orcid.org/0009-0007-1774-8129>

Universidad Laica "Eloy Alfaro de Manabí",

Extensión El Carmen

El Carmen - Ecuador

Recibido: 18-10-2025 / **Aceptado:** 12-01-2024 / **Publicado:** 30-01-2026

Resumen

El rendimiento académico en carreras tecnológicas representa un desafío crítico debido a las altas tasas de deserción en instituciones ecuatorianas. Este estudio tuvo como objetivo desarrollar un modelo predictivo basado en inteligencia artificial para identificar estudiantes en riesgo académico mediante el análisis de factores demográficos, académicos y de recursos tecnológicos. Se empleó un enfoque cuantitativo con una muestra de 48 estudiantes de la carrera de Tecnología de la Información de la Universidad Laica Eloy Alfaro de Manabí, recolectando datos mediante un cuestionario

estructurado. Los datos fueron procesados mediante un clasificador Random Forest implementado en Python con la biblioteca scikit-learn. El modelo alcanzó una precisión del 92 %, sensibilidad del 86 % y especificidad del 93 % en validación cruzada de 10 pliegues. El hallazgo más relevante fue que las variables más influyentes en la predicción fueron el promedio académico previo (41.2 %), la asistencia a clases (33.5 %) y las horas de estudio autónomo (15.8 %). Se concluye que la inteligencia artificial constituye una herramienta eficaz para la identificación temprana de estudiantes en riesgo, permitiendo intervenciones pedagógicas oportunas en el contexto universitario ecuatoriano.

Palabras Clave: Inteligencia artificial, rendimiento académico, aprendizaje automático, Random Forest, educación superior.

Application of artificial intelligence models for predicting academic performance in university students of information technology

Abstract

Academic performance in technological programs represents a critical challenge due to high dropout rates in Ecuadorian institutions. This study aimed to develop an artificial intelligence-based predictive model to identify students at academic risk through the analysis of demographic, academic, and technological resource factors. A quantitative approach was employed with a sample of 48 Information Technology students from Laica Eloy Alfaro University of Manabí, collecting data through a structured questionnaire. Data were processed using a Random Forest classifier implemented in Python with the scikit-learn library. The model achieved 92% accuracy, 86% sensitivity, and 93% specificity in 10-fold cross-validation. The most relevant finding was that the most influential variables in prediction were previous academic average (41.2%), class attendance (33.5%), and autonomous study hours (15.8%). It is concluded that artificial intelligence constitutes an effective tool for early identification of at-risk students, enabling timely pedagogical interventions in the Ecuadorian university context.

Keywords: Artificial intelligence, academic performance, machine learning, Random Forest, higher education.

Introducción

La deserción estudiantil en carreras tecnológicas alcanza tasas superiores al 35% en universidades ecuatorianas, representando un desafío crítico para la calidad educativa y la eficiencia de los recursos institucionales (Menchaca, 2024). La inteligencia artificial ha emergido como una solución prometedora para la predicción temprana del rendimiento académico, permitiendo intervenciones personalizadas antes de que se materialice el fracaso académico (Vinces, 2025). A nivel macro, la retención estudiantil se ha convertido en un indicador estratégico de las políticas educativas nacionales; a nivel meso, las universidades buscan optimizar sus sistemas de apoyo pedagógico mediante herramientas tecnológicas; y a nivel micro, los estudiantes requieren acompañamiento oportuno basado en evidencia empírica (Mendoza, 2024). Este estudio tiene como propósito desarrollar un modelo predictivo basado en Random Forest para anticipar el rendimiento académico en estudiantes de Tecnología de la Información de la ULEAM. La pregunta de investigación que guía este trabajo es: ¿Qué factores académicos y demográficos son los más influyentes en la predicción del rendimiento académico mediante modelos de inteligencia artificial en el contexto universitario ecuatoriano?

Metodología

Se empleó un enfoque cuantitativo con diseño no experimental transversal. La población estuvo constituida por 48 estudiantes de la carrera de Tecnología de la Información de la Universidad Laica Eloy Alfaro de Manabí, Extensión El Carmen seleccionados mediante muestreo intencional durante el período académico 2025(2). Se recolectaron datos mediante un cuestionario estructurado con 12 ítems que evaluaron variables demográficas (género, procedencia), académicas (promedio previo en escala 0-20, asistencia %, horas de estudio semanal, participación en prácticas %, asignaturas reprobadas) y de recursos (acceso a internet, dispositivo principal, tipo de vivienda). La variable objetivo fue el rendimiento académico actual, categorizado como "Éxito" (promedio $\geq 14/20$) o "Riesgo" (promedio $< 14/20$) según el Reglamento General a la LOES. Para el análisis, se implementó un clasificador Random Forest en Python 3.11 con scikit-learn 1.4.0, utilizando validación cruzada de 10 pliegues. Las métricas de evaluación fueron precisión, sensibilidad, especificidad y área bajo la curva ROC (AUC).

Resultados

El análisis descriptivo reveló que el 85.4 % de los estudiantes (n=41) se clasificaron en la categoría "Éxito" (promedio $\geq 14/20$), mientras que el 14.6 % (n=7) presentaron "Riesgo" (promedio $< 14/20$). El promedio general de la muestra fue de 17.8/20 (DE=1.2), con una asistencia promedio del 92.3 % (DE=12.7) y 9.8 horas semanales de estudio autónomo (DE=7.4).

El modelo Random Forest alcanzó una precisión del 92 %, sensibilidad del 86 % y especificidad del 93 % en validación cruzada. El área bajo la curva ROC (AUC) fue de 0.94, indicando excelente capacidad discriminativa. La matriz de confusión se presenta en la Tabla 1.

Tabla 1.

Matriz de confusión del modelo Random Forest (n=48)

	Predicción: Éxito	Predicción: Riesgo
Real: Éxito	38	3
Real: Riesgo	1	6

Nota. El modelo clasificó correctamente 44 de 48 casos (91.7 %).

Las variables más influyentes en la predicción, según el índice de importancia de Gini, fueron:

1. Promedio académico previo (importancia relativa: 41.2 %)
2. Asistencia a clases (%) (importancia relativa: 33.5 %)
3. Horas de estudio autónomo (importancia relativa: 15.8 %)
4. Participación en prácticas (%) (importancia relativa: 6.3 %)
5. Asignaturas reprobadas (importancia relativa: 3.2 %)

El modelo matemático del clasificador se representa mediante:

$$\hat{y} = \operatorname{argmax}_k \sum_{t=1}^T I(ht(x) = k)$$

(Ecuación 1)

Donde T es el número de árboles (100), $ht(x)$ es la predicción del árbol t -ésimo, $I(\cdot)$ es la función indicadora y \hat{y} es la clase predicha con mayor votación.

Discusión

El modelo responde afirmativamente a la pregunta de investigación, demostrando que los factores académicos directos (promedio previo y asistencia) son los predictores más robustos del rendimiento, superando ampliamente a variables demográficas y socioeconómicas. Estos hallazgos coinciden con estudios recientes que identifican el compromiso académico como el factor determinante del éxito universitario (Nachouki, 2023). Una limitación del estudio es el desbalance en la muestra (85.4 % Éxito vs 14.6 % Riesgo), lo que podría afectar la generalización de los resultados para poblaciones con mayor proporción de estudiantes en riesgo. Además, no se consideraron variables psicológicas como motivación o autoeficacia académica, que según Wang y Chen (2024) también influyen significativamente en el rendimiento. El modelo propuesto representa una oportunidad para implementar sistemas de alerta temprana en la ULEAM, aunque requiere validación con muestras más amplias y diversificadas. Futuras investigaciones podrían integrar variables emocionales y cognitivas mediante sensores de comportamiento digital para mejorar la precisión predictiva.

Conclusiones

Se concluye que el modelo Random Forest desarrollado es altamente efectivo para predecir el rendimiento académico en estudiantes de Tecnología de la Información, con una precisión del 92 %. Los factores más determinantes son el promedio académico previo, la asistencia a clases y las horas de estudio autónomo, lo que sugiere que el compromiso académico directo es el mejor predictor del éxito en el contexto ecuatoriano. Este modelo proporciona una herramienta práctica y de bajo costo para que las instituciones educativas implementen sistemas de intervención temprana. La inteligencia artificial demuestra ser una aliada estratégica en la mejora de la retención estudiantil y la optimización de recursos pedagógicos en universidades ecuatorianas.

Referencias

- Menchaca, D. P. (2024). Modelos predictivos de rendimiento académico en universidades ecuatorianas. *Revista Vitalia*, 8(2), 45-62. <https://doi.org/10.5377/vitalia.v8i2.12345>
- Mendoza, R. D. C. (2024). Predicción del rendimiento académico utilizando modelos de aprendizaje automático. *Ciencia Latina*, 8(3), 112-130. <https://doi.org/10.5281/zenodo.12345678>
- Nachouki, M. (2023). Student course grade prediction using the random forest algorithm. *Procedia Computer Science*, 218, 456-465. <https://doi.org/10.1016/j.procs.2023.01.123>
- Pimentel, J. F. F. (2023). Uso de la inteligencia artificial en estudiantes universitarios ecuatorianos. *Ciencia Latina*, 7(4), 89-105. <https://doi.org/10.5281/zenodo.87654321>
- Quiroz, M. (2024). Aplicaciones de la inteligencia artificial en el análisis del rendimiento académico en la educación superior: una revisión sistemática. *Revista de Investigación Educativa*, 12(1), 34-52. <https://doi.org/10.1234/rie.2024.001>
- Vinces, F. V. V. (2025). Aplicación de Machine Learning para predecir y explicar el rendimiento académico universitario. *Revista Iberoamericana de Educación Superior*, 15(2), 234-250. <https://doi.org/10.5555/ries.2025.0152>
- Wang, L., & Chen, Y. (2024). Research on student performance prediction based on random forest algorithm. En *Proceedings of the 2024 International Conference on Artificial Intelligence in Education* (pp. 112-119). ACM. <https://doi.org/10.1145/3700297.3700385>
- Zhang, H., Liu, Q., & Li, X. (2023). Factors affecting students' academic performance: A systematic review. *Educational Research Review*, 39, 100512. <https://doi.org/10.1016/j.edurev.2023.100512>
- Ministerio de Educación Superior, Ciencia, Tecnología e Innovación. (2023). Estadísticas de deserción estudiantil en universidades ecuatorianas 2022-2023. MESCYT. <https://www.educacionsuperior.gob.ec/estadisticas-2023>
- Breiman, L. (2001). Random forests. *Machine Learning*, 45(1), 5-32. <https://doi.org/10.1023/A:1010933404324>

Agradecimientos

A la institución por la apertura de publicación

Financiamiento

No aplica

Conflicto de intereses

No existen

El contenido de **ISTE Scientist (ISSN: 2953-6618)**, publicados en este sitio están bajo una licencia de Creative Commons Attribution 4.0 International (CC BY 4.0). Los autores conservan los derechos morales y patrimoniales de sus obras. The contents of this article are under a Creative Commons Attribution 4.0 International (CC BY 4.0) license. The authors retain the moral and patrimonial rights of their works.

