

Research Article

Inteligencia artificial para la predicción temprana de la deserción estudiantil en educación superior: revisión sistemática

Self-Supervised Learning Process for Early Academic Monitoring Using Artificial Intelligence: A Systematic Review



Cedeño-Salazar, Bladimir ¹



<https://orcid.org/0000-0001-5397-7067>



bcedeno@uteq.edu.ec



Ecuador, Quevedo, Universidad Técnica Estatal de Quevedo



Mendoza-Vargas, Emma Yolanda ³



<https://orcid.org/0000-0002-0220-4328>



emendoza@uteq.edu.ec



Ecuador, Quevedo, Universidad Técnica Estatal de Quevedo



Díaz-Macías Efraín ²



<https://orcid.org/0000-0003-4087-029X>



efraindiaz@uteq.edu.ec



Ecuador, Quevedo, Universidad Técnica Estatal de Quevedo



Escobar-Terán, Harold ⁴



<https://orcid.org/0000-0001-9165-6627>



hescobar@uteq.edu.ec



Ecuador, Quevedo, Universidad Técnica Estatal de Quevedo

Autor de correspondencia ¹



DOI / URL: <https://doi.org/10.55813/gaea/jessr/v6/n3/274>

Resumen: La deserción estudiantil constituye un desafío prioritario para la educación superior debido a sus efectos académicos, sociales e institucionales. Este estudio analiza dicha problemática en la Universidad Técnica Estatal de Quevedo (UTEQ) mediante una revisión sistemática de la literatura orientada a identificar los principales factores asociados y examinar el potencial de la inteligencia artificial para fortalecer el seguimiento académico. La metodología se desarrolló conforme a los lineamientos PRISMA, garantizando transparencia en la búsqueda, selección y evaluación de los estudios. De más de 500 publicaciones inicialmente identificadas, se incluyeron 20 por su pertinencia temática y metodológica. Los resultados evidenciaron que los factores socioeconómicos y académicos fueron los más influyentes, con una representación del 30 % cada uno. Asimismo, Colombia concentró la mayor producción científica, seguida de Perú, México y Chile. Se propone incorporar técnicas de aprendizaje automático para detectar tempranamente estudiantes en riesgo, personalizar intervenciones y mejorar la retención, el desempeño y la continuidad educativa en la UTEQ.

Palabras clave: abandono universitario, aprendizaje automático, minería de datos, analítica educativa, retención académica.



Check for
updates

Receptado: 06/Jun/2026

Aceptado: 03/Jul/2026

Publicado: 31/Jul/2026

Cita: Cedeño-Salazar, B., Díaz-Macías, E., Mendoza-Vargas, E. Y., & Escobar-Terán, H. (2026). Inteligencia artificial para la predicción temprana de la deserción estudiantil en educación superior: revisión sistemática. *Journal of Economic and Social Science Research*, 6(3), 193-210. <https://doi.org/10.55813/gaea/jessr/v6/n3/274>

Journal of Economic and Social Science Research (JESSR)
<https://economicsocialresearch.com>
jessr@editorialgrupo-aea.com
info@editorialgrupo-aea.com

Nota del editor: Editorial Grupo AEA se mantiene neutral con respecto a las reclamaciones legales resultantes de contenido publicado. La responsabilidad de información publicada recae enteramente en los autores.

© 2026. Este artículo es un documento de acceso abierto distribuido bajo los términos y condiciones de la **Licencia Creative Commons, Atribución-NoComercial 4.0 Internacional**.



Abstract:

Abstract: Student dropout is a key challenge for higher education due to its academic, social, and institutional impacts. This study analyzes this issue at the Quevedo State Technical University (UTEQ) through a systematic literature review aimed at identifying the main associated factors and examining the potential of artificial intelligence to strengthen academic monitoring. The methodology was developed in accordance with PRISMA guidelines, ensuring transparency in the search, selection, and evaluation of studies. Of the more than 500 publications initially identified, 20 were included based on their thematic and methodological relevance. The results showed that socioeconomic and academic factors were the most influential, each accounting for 30% of the total. Furthermore, Colombia accounted for the largest volume of scientific output, followed by Peru, Mexico, and Chile. The study proposes incorporating machine learning techniques to detect at-risk students early, personalize interventions, and improve retention, performance, and educational continuity at UTEQ.

Keywords: college dropout, machine learning, data mining, educational analytics, academic retention.

1. Introducción

La deserción estudiantil constituye la interrupción temporal o definitiva de la trayectoria formativa antes de completar un programa académico. Su aparición no responde a una causa aislada, sino a la interacción de condiciones personales, familiares, económicas, académicas e institucionales que afectan la permanencia universitaria. El rendimiento previo, la integración con el entorno educativo y la disponibilidad de recursos figuran entre los factores explicativos más frecuentes (Parra-Sánchez et al., 2023). Por ello, su análisis requiere una perspectiva integral que permita reconocer riesgos, comprender sus relaciones y orientar intervenciones oportunas desde la gestión institucional (Sanhueza et al., 2021).

En América Latina, la deserción universitaria representa un problema persistente por sus consecuencias sobre los estudiantes, las instituciones y los sistemas de educación superior. En Chile se han identificado diferencias entre universidades públicas relacionadas con la gestión institucional y las condiciones de permanencia (Sanhueza et al., 2021). En otros contextos regionales, el análisis de datos académicos ha permitido establecer asociaciones entre desempeño, antecedentes estudiantiles y probabilidad de abandono (Miranda & Guzmán, 2017). Estas evidencias muestran que el fenómeno adopta características particulares según el entorno, por lo que las estrategias preventivas deben responder a las condiciones de cada institución (Casanova-Villalba et al., 2023).

En Ecuador, la reducción de la deserción fue incorporada como prioridad en el Plan de Creación de Oportunidades 2021-2025, que estableció disminuir la tasa de abandono durante el primer año de educación superior del 21,84 % al 19,89 % (Secretaría Nacional de Planificación, 2021). El cumplimiento de esta meta exige mecanismos capaces de identificar dificultades antes de que ocasionen la desvinculación académica. Los modelos predictivos muestran mayor utilidad cuando integran variables socioeconómicas, demográficas y de rendimiento en un mismo proceso analítico (Villarreal-Torres et al., 2024).

En la Universidad Técnica Estatal de Quevedo, el seguimiento de los estudiantes de grado se desarrolla mediante tutorías, encuestas socioeconómicas, entrevistas y revisión periódica de calificaciones. Aunque estas acciones permiten reconocer determinadas dificultades, su aplicación puede depender de valoraciones manuales y de información examinada cuando el riesgo ya se ha manifestado. La minería de datos facilita la identificación de patrones asociados con la permanencia y el abandono a partir del desempeño académico (Pérez-Gutiérrez, 2020). De igual manera, la clasificación automatizada permite agrupar a los estudiantes según niveles de riesgo y apoyar la priorización del acompañamiento institucional (Eckert & Suénaga, 2015).

La inteligencia artificial ofrece alternativas para fortalecer el seguimiento académico mediante el procesamiento de numerosos datos y la generación de predicciones individualizadas. Los algoritmos de aprendizaje automático permiten reconocer relaciones que podrían pasar inadvertidas mediante procedimientos convencionales y anticipar casos de abandono (Kuz & Morales, 2023). Las revisiones disponibles destacan el empleo de técnicas de clasificación, minería de datos y modelos predictivos para analizar el rendimiento universitario (Jimbo-Santana et al., 2023). Sin embargo, su utilidad depende de la calidad de los datos, la selección de atributos y la incorporación de los resultados en la gestión académica (Agrusti et al., 2019).

A partir de esta problemática, el estudio tuvo como objetivo analizar la aplicación de la inteligencia artificial en la predicción temprana de la deserción estudiantil en educación superior y valorar su contribución al seguimiento académico de las carreras de grado de la Universidad Técnica Estatal de Quevedo. Para ello, se desarrolló una revisión sistemática de publicaciones comprendidas entre 2015 y 2024 en IEEE, ScienceDirect y Scopus. El análisis permitió identificar los factores predictivos más frecuentes, las técnicas empleadas y las tendencias investigativas, con el propósito de formular orientaciones para un sistema institucional de seguimiento oportuno, personalizado y eficiente (Piedra-Castro et al., 2024).

2. Materiales y métodos

Diseño del estudio

Se desarrolló una investigación de enfoque mixto y alcance descriptivo, con predominio documental. El procedimiento integró dos componentes complementarios:

una revisión sistemática de la literatura sobre la aplicación de inteligencia artificial en la predicción de la deserción estudiantil y una fase de entrevistas dirigida a tutores académicos y personal técnico-administrativo de la Universidad Técnica Estatal de Quevedo. La combinación de ambas fuentes permitió examinar la evidencia científica disponible y contrastarla con las condiciones institucionales del seguimiento académico. No se efectuó una intervención experimental ni se implementó un modelo predictivo en la universidad; el estudio se orientó al análisis de métodos, factores y posibilidades de aplicación.

La revisión sistemática se desarrolló conforme a las directrices de la declaración PRISMA 2020, mediante las etapas de identificación, cribado, recuperación, elegibilidad e inclusión de publicaciones (Page et al., 2021). Para estructurar la pregunta principal y delimitar los componentes de la búsqueda se empleó la estrategia PICO, integrada por población o problema, intervención, comparación y resultado (Santos et al., 2007).

Formulación de la pregunta mediante PICO

En la fase inicial se definieron los componentes de la estrategia PICO. La población correspondió a los estudiantes de las carreras de grado de la Universidad Técnica Estatal de Quevedo; la intervención comprendió el seguimiento académico apoyado mediante inteligencia artificial; la comparación se estableció con los métodos convencionales de seguimiento; y el resultado se relacionó con la predicción temprana de la deserción estudiantil.

Tabla 1
Componentes de la estrategia PICO utilizados en la investigación

Inicial	Identificador	Elemento de la investigación
P	Población o problema	Estudiantes de carreras de grado de la Universidad Técnica Estatal de Quevedo
I	Intervención o exposición	Seguimiento académico mediante inteligencia artificial
C	Comparación	Métodos convencionales de seguimiento académico
O	Resultados	Predicción de la deserción estudiantil

Nota: (Autores, 2026).

A partir de estos componentes se formuló la pregunta principal: ¿Cómo influyó el seguimiento académico mediante inteligencia artificial en la predicción de la deserción estudiantil en las carreras de grado de la Universidad Técnica Estatal de Quevedo, en comparación con los métodos convencionales?

La pregunta principal se desagregó en cuatro interrogantes secundarias orientadas a examinar la evolución temporal de las investigaciones, su distribución geográfica, los factores asociados con la deserción y los métodos de inteligencia artificial empleados. Estas preguntas dirigieron la búsqueda, la extracción de información y la organización de los resultados.

Tabla 2
Pregunta principal y preguntas secundarias de la revisión sistemática

Nro.	Preguntas de investigación
RQ1.	¿Cómo influye el seguimiento académico mediante inteligencia artificial en la predicción de la deserción estudiantil en las carreras de grado de la Universidad Técnica Estatal de Quevedo en comparación con métodos convencionales?
RQ1.1.	¿Cómo han evolucionado los métodos de inteligencia artificial aplicados en la deserción estudiantil en la educación universitaria?
RQ1.2.	¿En qué países latinoamericanos se han desarrollado estas investigaciones sobre aplicación de la inteligencia artificial en la deserción estudiantil en la educación universitaria?
RQ1.3.	¿Qué factores han influido en la deserción estudiantil en la educación universitaria?
RQ1.4.	¿Cuáles son los métodos de inteligencia artificial más utilizados para integrarlos en la predicción de la deserción estudiantil?

Nota: (Autores, 2026).

Fuentes de información y periodo de búsqueda

La búsqueda bibliográfica se realizó en IEEE Xplore, ScienceDirect y Scopus, seleccionadas por su cobertura de publicaciones relacionadas con ciencias computacionales, inteligencia artificial, educación superior y análisis predictivo. Se consideraron trabajos publicados entre 2015 y 2024, de acuerdo con la delimitación temporal establecida para la revisión.

La consulta se orientó hacia investigaciones sobre inteligencia artificial, aprendizaje automático, aprendizaje profundo, minería de datos, rendimiento académico, seguimiento estudiantil y predicción de la deserción en instituciones de educación superior. La búsqueda inicial recuperó 500 registros: 105 procedentes de IEEE Xplore, 178 de ScienceDirect y 217 de Scopus.

Estrategia de búsqueda

Los términos se definieron a partir de los componentes PICO y las preguntas de investigación. Los sinónimos de cada bloque temático se combinaron mediante el operador OR, mientras que los bloques se relacionaron mediante AND. La búsqueda se adaptó a IEEE Xplore, ScienceDirect y Scopus; en esta última se consultaron los campos de título, resumen y palabras clave.

Tabla 3
Términos y cadena de búsqueda utilizados

Bloque	Términos
Inteligencia artificial	"Artificial Intelligence" OR "Machine Learning" OR "Deep Learning" OR "AI"
Deserción estudiantil	"Student Dropout" OR "Academic Attrition" OR "University Withdrawal" OR "Retention"
Educación superior	"Higher Education" OR "University" OR "Tertiary Education"
Enfoque analítico	"Evolution" OR "Methods" OR "Techniques" OR "Models" OR "Factors" OR "Prediction"
Cadena general	("Artificial Intelligence" OR "Machine Learning" OR "Deep Learning" OR "AI") AND ("Student Dropout" OR "Academic Attrition" OR "University Withdrawal" OR "Retention") AND ("Higher Education" OR "University" OR "Tertiary Education") AND ("Evolution" OR "Methods" OR "Techniques" OR "Models" OR "Factors" OR "Prediction")

Nota: (Autores, 2026).

Criterios de elegibilidad

Los criterios se establecieron antes de seleccionar las publicaciones y se aplicaron durante el cribado de títulos, resúmenes y textos completos.

Tabla 4
Criterios de inclusión y exclusión

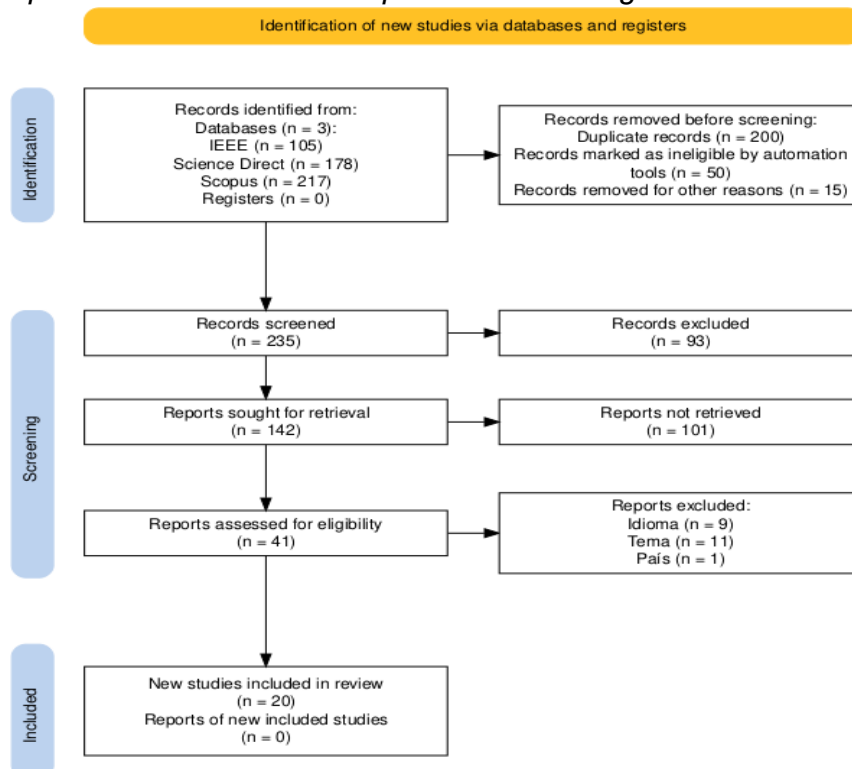
Inclusión	Exclusión
Publicaciones comprendidas entre 2015 y 2024.	Registros duplicados.
Estudios desarrollados en educación superior.	Publicaciones fuera del periodo establecido.
Investigaciones realizadas en países latinoamericanos.	Estudios ajenos a la educación superior o desarrollados fuera de Latinoamérica.
Trabajos sobre inteligencia artificial, aprendizaje automático, aprendizaje profundo o minería de datos.	Documentos sin métodos o resultados vinculados con inteligencia artificial.
Estudios relacionados con predicción, clasificación o identificación del riesgo de deserción.	Publicaciones sin relación directa con la deserción estudiantil.
Investigaciones que presentaron factores o resultados asociados con el abandono universitario.	Documentos redactados en idiomas no considerados en el protocolo.

Nota: (Autores, 2026).

Selección de las publicaciones

La búsqueda recuperó 500 registros: 105 de IEEE Xplore, 178 de ScienceDirect y 217 de Scopus. Antes del cribado se eliminaron 200 duplicados, 50 registros considerados no elegibles mediante herramientas de automatización y 15 documentos por otros motivos, por lo que quedaron 235 publicaciones.

Durante el cribado se excluyeron 93 registros y se intentó recuperar el texto completo de 142 publicaciones; 101 no pudieron obtenerse. De los 41 documentos evaluados, se excluyeron nueve por el idioma, 11 por falta de relación temática y uno por corresponder a un país no latinoamericano. Finalmente, se incluyeron 20 estudios.

Figura 1*Diagrama del proceso de selección de publicaciones según PRISMA 2020*

Nota: (Autores, 2026).

Extracción y análisis de la información

De cada publicación se extrajeron el título, año, país, propósito, factores asociados con la deserción, técnicas de inteligencia artificial, algoritmos y principales resultados. La información se organizó en una matriz y se analizó mediante frecuencias absolutas, porcentajes y síntesis narrativa. No se realizó metaanálisis debido a la heterogeneidad de los diseños, variables y métricas reportadas.

Entrevistas a actores institucionales

Se entrevistó a 20 tutores académicos y dos integrantes del personal técnico-administrativo de la UTEQ. En los tutores se examinaron el seguimiento académico, las herramientas utilizadas, los indicadores de riesgo, las causas atribuidas al abandono y las estrategias de apoyo. En el personal técnico-administrativo se analizaron la infraestructura, la actualización de los datos y las funcionalidades del sistema institucional. Las respuestas se resumieron mediante frecuencias, porcentajes y categorías temáticas.

Triangulación de la información

Los hallazgos de la revisión sistemática se contrastaron con las respuestas de los tutores y del personal técnico-administrativo mediante triangulación de fuentes (Okuda & Gómez-Restrepo, 2005). Se examinaron coincidencias y diferencias relacionadas con los factores de deserción, el acceso a la información, las limitaciones del seguimiento y las posibilidades de incorporar herramientas predictivas.

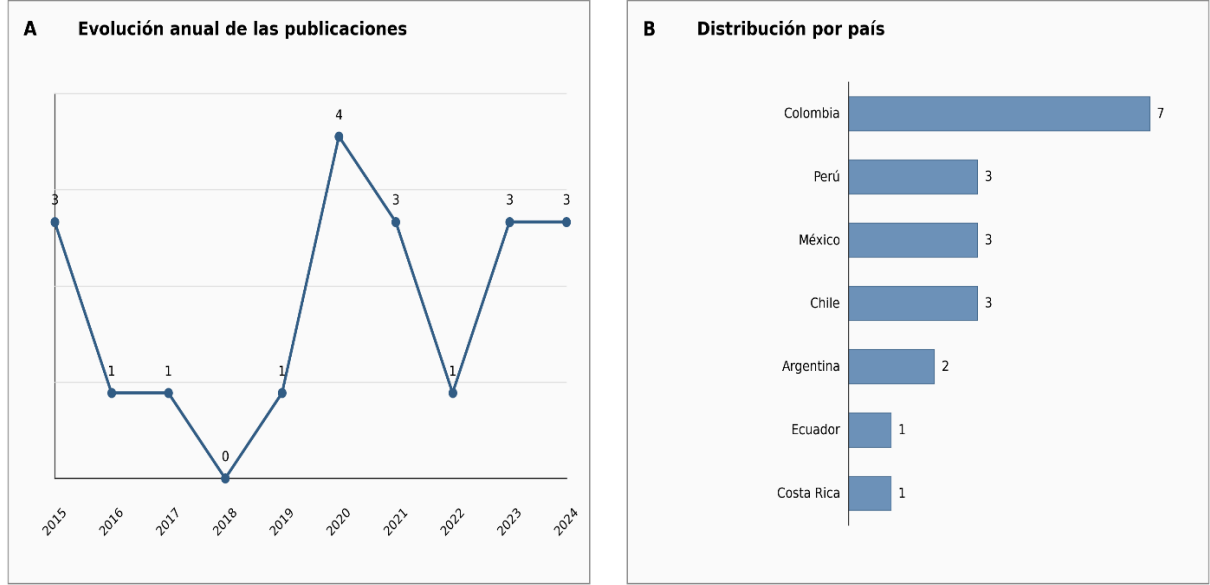
3. Resultados

3.1. Evolución temporal y distribución geográfica de los estudios

La producción científica incluida en la revisión presentó una distribución irregular entre 2015 y 2024. En 2015 se registraron tres publicaciones; en 2016 y 2017 se identificó una por año; y en 2018 no se incluyeron estudios. La mayor frecuencia se alcanzó en 2020, con cuatro artículos. Posteriormente, se registraron tres publicaciones en 2021, una en 2022 y tres tanto en 2023 como en 2024. En conjunto, 14 de los 20 estudios seleccionados se publicaron entre 2020 y 2024, lo que representó el 70 % de la muestra analizada.

La distribución geográfica mostró una concentración de la producción en Colombia, país del que procedieron siete investigaciones, equivalentes al 35 % del total. Perú, México y Chile aportaron tres publicaciones cada uno, con una participación individual del 15 %. Argentina registró dos estudios, mientras que Ecuador y Costa Rica presentaron uno respectivamente. De esta manera, Colombia, Perú, México y Chile reunieron 16 publicaciones y concentraron el 80 % de los trabajos incluidos, mientras que los tres países restantes representaron conjuntamente el 20 % (Figura 2).

Figura 2
Evolución temporal y distribución geográfica de las publicaciones incluidas



Nota: El panel A presenta el número de publicaciones por año y el panel B su distribución por país. Total de estudios incluidos: n = 20 (Autores, 2026).

3.2. Factores asociados y métodos de inteligencia artificial

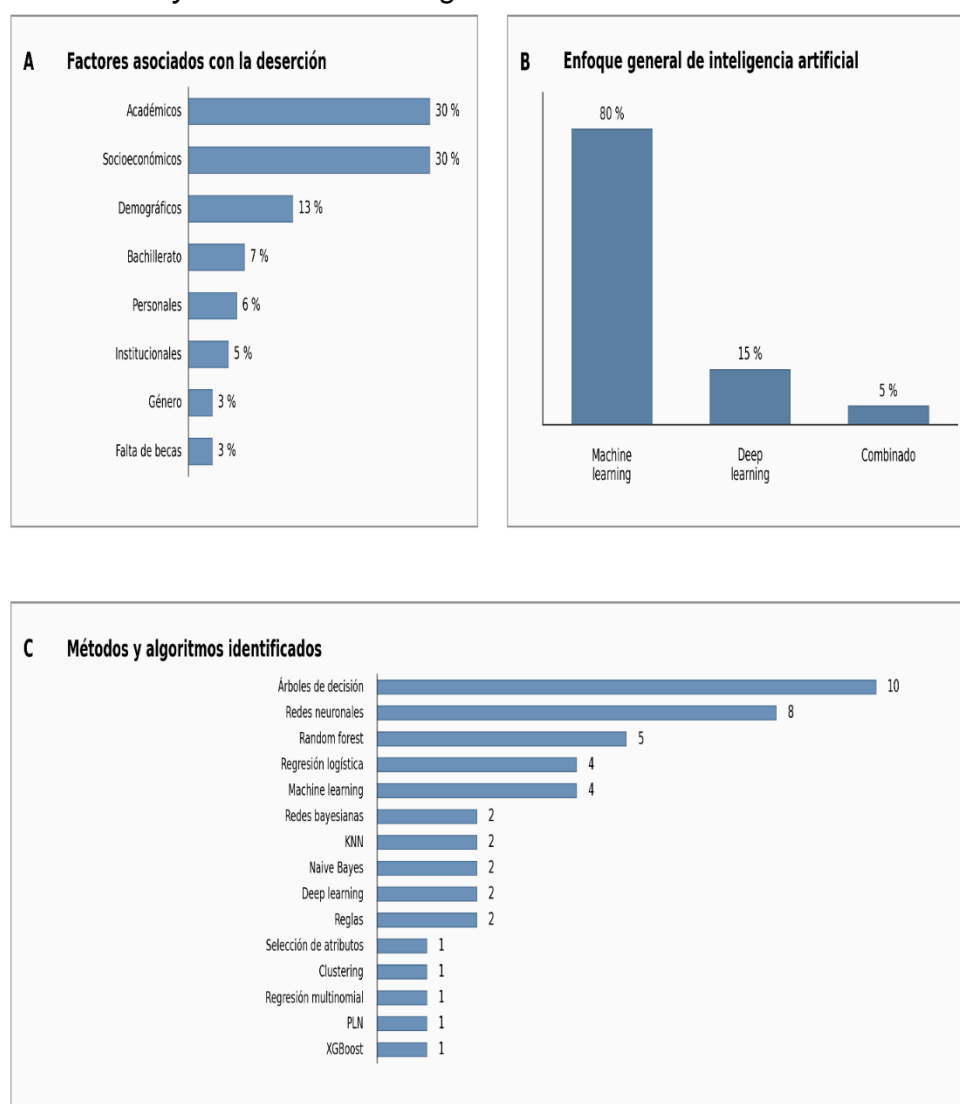
Los factores académicos y socioeconómicos fueron los componentes con mayor presencia en los estudios examinados y representaron el 30 % de las menciones cada uno. El artículo original señaló que ambas categorías estuvieron identificadas en 18 de las 20 publicaciones analizadas. Los factores demográficos alcanzaron el 13 %, seguidos por el nivel de bachillerato con el 7 %, los factores personales con el 6 % y

los institucionales con el 5 %. El género y la falta de becas presentaron las proporciones más bajas, con el 3 % respectivamente.

Los árboles de decisión constituyeron la técnica más utilizada, al aparecer en 10 estudios, seguidos por las redes neuronales en ocho y random forest en cinco. La regresión logística y las aplicaciones generales de machine learning se identificaron en cuatro publicaciones cada una. Las demás técnicas aparecieron entre uno y dos estudios. Al agrupar los trabajos por enfoque general, el 80 % empleó principalmente machine learning, el 15 % utilizó deep learning y el 5 % combinó ambos. La frecuencia de métodos superó el número de artículos porque varias investigaciones evaluaron más de un algoritmo.

Figura 3

Factores asociados y métodos de inteligencia artificial identificados en la revisión



Nota: El panel A muestra la proporción de menciones de los factores asociados; el panel B presenta el enfoque general de inteligencia artificial; y el panel C resume la frecuencia de los métodos y algoritmos reportados. Las categorías no fueron excluyentes (Autores, 2026).

Tabla 5

Métodos de inteligencia artificial y principales resultados de los estudios incluidos

N.º	Estudio	Métodos o técnicas	Resultado principal
1	Modelo de clasificación para la deserción estudiantil en universidades públicas del Perú (2024)	Machine learning supervisado mediante H2O.ai y gradient boosting.	El modelo permitió clasificar la deserción estudiantil mediante técnicas supervisadas.
2	Detección de alertas tempranas para prevenir la deserción estudiantil (2021)	Árboles de decisión, regresión logística y random forest.	Los modelos aportaron alertas tempranas y elementos para orientar políticas de retención.
3	Características que influyen en la deserción en una universidad latinoamericana (2023)	Árbol de decisión C4.5 mediante RapidMiner.	Se identificaron patrones vinculados con la deserción a partir de las características disponibles.
4	Análisis multinomial y predictivo de los factores asociados con la deserción (2019)	Regresión multinomial y random forest.	Random forest presentó el mejor desempeño y reportó un valor de 0,83 para reconocer al potencial desertor.
5	Análisis de la deserción mediante minería de datos (2017)	Redes bayesianas, árboles de decisión y redes neuronales.	Los tres algoritmos presentaron resultados semejantes para la predicción.
6	Modelo predictivo multivariable en programas virtuales de posgrado (2021)	Deep learning y procesamiento del lenguaje natural.	La combinación de ambas técnicas produjo resultados favorables para predecir el desempeño.
7	Modelo predictivo para detectar estudiantes con alto riesgo (2021)	Árboles de decisión, KNN, regresión logística y redes neuronales.	KNN alcanzó 88,844 % de acierto; el árbol de decisión obtuvo 88,4 %.
8	Patrones de deserción mediante minería de datos educativa (2020)	Selección de atributos y árbol de decisión C4.5.	El algoritmo C4.5 alcanzó una precisión del 92,6 %.
9	Análisis de deserción y permanencia mediante clasificación (2015)	Redes bayesianas, árboles de decisión y reglas.	Los criterios evaluados presentaron desempeños similares.
10	Caracterización de la deserción en educación superior (2015)	Clustering, árboles de decisión y reglas.	Se identificó un patrón asociado con promedio bajo y asignaturas reprobadas en los primeros semestres.
11	Ciencia de datos educativos y deserción universitaria en México (2023)	Árboles de decisión, regresión logística, XGBoost y redes neuronales.	XGBoost alcanzó una exactitud superior al 99 %; las demás técnicas no superaron el 94 %.
12	Modelo de aprendizaje profundo para universidades virtuales (2020)	Redes neuronales y deep learning.	Las redes neuronales fueron propuestas como herramienta para analizar la deserción.
13	Comparación de técnicas de minería de datos (2020)	Árboles de decisión, regresión logística, naive Bayes y random forest.	Random forest registró 0,91 en el tercer semestre y 0,97 en el último.

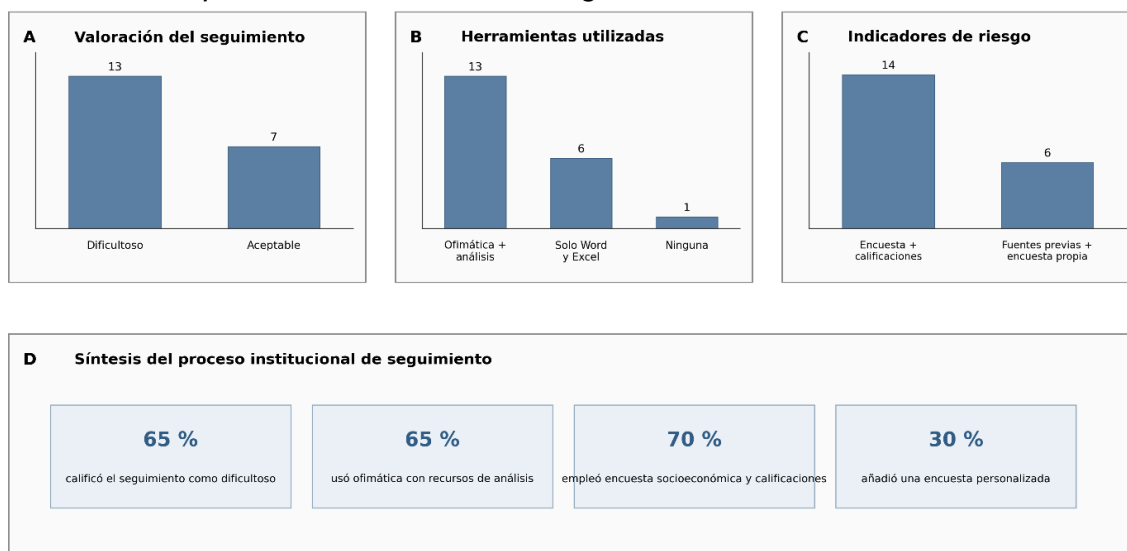
14	Minería de datos para detectar patrones de deserción (2015)	Árboles de decisión mediante Weka.	La técnica generó información para la toma de decisiones institucionales.
15	Variables asociadas con la deserción de estudiantes de primer año (2016)	Naive Bayes, random forest y redes neuronales.	Random forest alcanzó el mejor desempeño general, con 88,9 % de exactitud.
16	Inteligencia artificial para analizar el rendimiento académico (2023)	Redes neuronales y deep learning.	La revisión reportó resultados favorables de las redes neuronales para predecir el rendimiento.
17	Estrategias de retención estudiantil con inteligencia artificial (2024)	Machine learning supervisado.	Las técnicas supervisadas fueron empleadas para apoyar la predicción y la retención.
18	Inteligencia artificial para predecir el rendimiento y evitar el abandono (2024)	Redes neuronales y random forest.	Random forest fue señalado como una técnica eficaz para predecir el rendimiento.
19	Agrupación de estudiantes según riesgo de abandono (2022)	Machine learning supervisado y agrupación por riesgo.	La clasificación por niveles de riesgo facilitó la organización de acciones de acompañamiento.
20	Predicción temprana de la deserción en cursos profesionales en línea (2020)	KNN, machine learning y redes neuronales.	KNN presentó el mejor desempeño entre las técnicas comparadas.

Nota: (Autores, 2026).

3.3. Condiciones del seguimiento académico institucional

De los 20 tutores entrevistados, 13 calificaron el proceso de seguimiento académico como dificultoso, lo que representó el 65 % de los participantes; los siete restantes lo consideraron aceptable. En cuanto a los recursos utilizados, 13 tutores combinaron programas ofimáticos con herramientas adicionales para el procesamiento estadístico, seis trabajaron únicamente con Word y Excel y uno manifestó no utilizar herramientas tecnológicas. Por tanto, aunque 19 participantes emplearon algún recurso digital, el seguimiento se apoyó principalmente en aplicaciones ofimáticas y procedimientos gestionados de manera individual.

Respecto de los indicadores utilizados, 14 tutores señalaron que reconocían a los estudiantes en riesgo mediante la encuesta socioeconómica institucional y los registros de calificaciones de los dos parciales del semestre. Los seis participantes restantes incorporaron, además, una encuesta personalizada aplicada mediante herramientas digitales. La información evidenció que todos los tutores consideraban datos socioeconómicos y académicos, pero solo el 30 % complementaba esas fuentes con información adicional. En las respuestas registradas no se identificó el uso de un sistema institucional que integrara automáticamente los indicadores y generara niveles de riesgo.

Figura 4*Condiciones del proceso institucional de seguimiento académico*

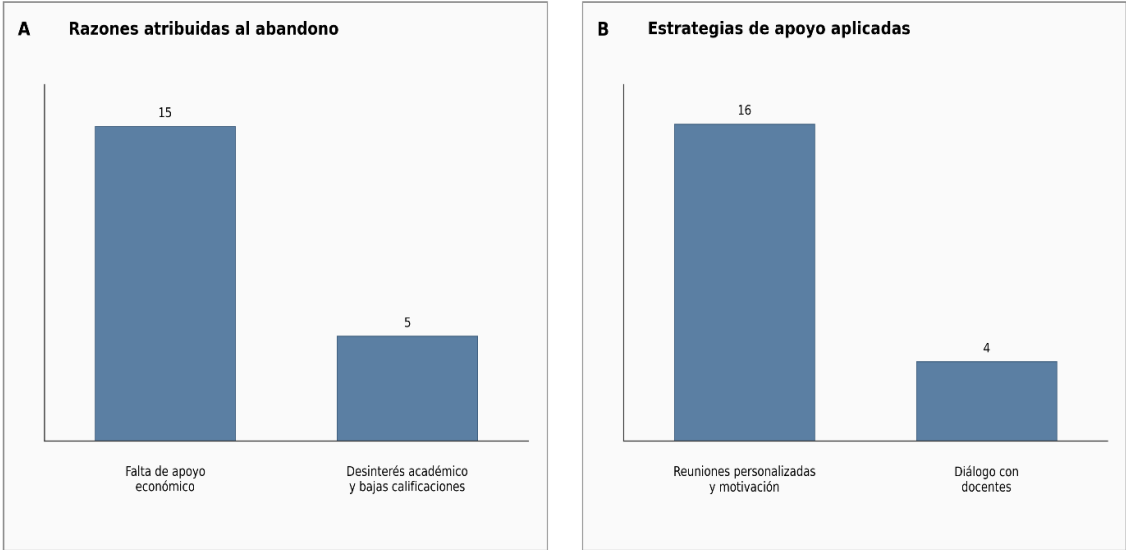
Nota: Los paneles A, B y C presentan la valoración del seguimiento, las herramientas utilizadas y los indicadores empleados. El panel D sintetiza las proporciones principales. Total de tutores: n = 20 (Autores, 2026).

3.4. Razones atribuidas al abandono y estrategias de apoyo

Quince de los 20 tutores entrevistados identificaron la falta de apoyo económico como la principal razón del abandono estudiantil, proporción equivalente al 75 %. Los cinco participantes restantes, correspondientes al 25 %, atribuyeron la interrupción de los estudios al desinterés académico y a la reprobación de asignaturas por bajas calificaciones. Las respuestas concentraron las causas percibidas en dos dimensiones: las condiciones económicas y el desempeño académico. Estos datos representaron valoraciones derivadas de la experiencia tutorial y no una medición directa de las causas entre la totalidad de estudiantes de la institución.

En relación con las acciones desarrolladas, 16 tutores, equivalentes al 80 %, indicaron que realizaban reuniones personalizadas y actividades de motivación con los estudiantes en riesgo. Los cuatro participantes restantes, correspondientes al 20 %, señalaron que recurrían al diálogo con los docentes de las asignaturas para mediar en los problemas académicos. En consecuencia, las estrategias se concentraron en el acompañamiento directo, la comunicación y la mediación docente. Las respuestas no registraron mecanismos automatizados que activaran intervenciones a partir de alertas o clasificaciones predictivas.

Figura 5
Razones atribuidas al abandono y estrategias de apoyo aplicadas por los tutores



Nota: El panel A presenta las razones atribuidas al abandono y el panel B las estrategias de apoyo empleadas. Total de tutores: n = 20 (Autores, 2026).

3.5. Integración de los resultados documentales e institucionales

Las entrevistas a dos integrantes del personal técnico-administrativo indicaron que la información académica y socioeconómica requería procesos de limpieza y mejora antes de utilizarse con fines analíticos. También se señaló que no existía una normativa específica para actualizar la información socioeconómica y que el sistema institucional permitía consultar las calificaciones, pero no generaba alertas asociadas con el desempeño. Estos hallazgos se integraron con la revisión sistemática y las respuestas de los tutores para identificar coincidencias y diferencias en los factores, los datos empleados, las herramientas disponibles y las formas de intervención.

Tabla 7
Integración de los resultados según la fuente de información

Dimensión	Revisión sistemática	Tutores académicos	Personal técnico-administrativo	Síntesis
Factores asociados	Predominaron los factores académicos y socioeconómicos.	El 75 % señaló la falta de apoyo económico y el 25 % el desinterés académico y las bajas calificaciones.	No se incorporó una clasificación causal adicional.	Coincidencia entre la evidencia revisada y las percepciones de los tutores.
Datos e indicadores	Los modelos combinaron variables académicas, socioeconómicas, demográficas y personales.	El 70 % utilizó encuestas socioeconómicas y calificaciones; el 30 % añadió una encuesta propia.	Se reconoció la necesidad de limpiar y mejorar los datos.	La información institucional fue más limitada y fragmentada que la utilizada en los modelos revisados.

Dimensión	Revisión sistemática	Tutores académicos	Personal técnico-administrativo	Síntesis
Herramientas	Se identificaron árboles de decisión, redes neuronales, random forest, regresión logística y otros algoritmos.	El seguimiento se apoyó principalmente en programas ofimáticos y herramientas estadísticas.	El sistema permitía consultar calificaciones, pero no generaba alertas.	Se observó una diferencia entre las capacidades predictivas descritas y las herramientas institucionales disponibles.
Intervención	Los estudios propusieron alertas tempranas, clasificación del riesgo y apoyo a decisiones de retención.	El 80 % realizó reuniones y motivación; el 20 % recurrió a la mediación con docentes.	No se reportó un mecanismo automatizado para activar intervenciones.	Las acciones institucionales fueron directas y manuales, sin activación predictiva.
Condiciones institucionales	La utilidad de los modelos dependió de datos suficientes, actualizados y organizados.	El 65 % consideró difícil el seguimiento por limitaciones de información.	No existía normativa específica para actualizar los datos socioeconómicos.	La calidad y actualización de los datos constituyeron condiciones críticas para una futura aplicación.

Nota: (Autores, 2026).

4. Discusión

La concentración del 70 % de las publicaciones entre 2020 y 2024 evidenció un interés reciente por aplicar inteligencia artificial a la predicción de la deserción universitaria. Esta tendencia coincidió con la expansión de técnicas de aprendizaje automático destinadas a analizar el rendimiento, la permanencia y el riesgo de abandono estudiantil (Cruz et al., 2022; Jimbo-Santana et al., 2023). Sin embargo, la producción se concentró principalmente en Colombia, Perú, México y Chile, mientras que Ecuador presentó una participación limitada. Esta distribución mostró que el desarrollo regional no fue homogéneo y que los modelos predictivos deberían adaptarse a las condiciones académicas, sociales y administrativas de cada institución antes de su implementación (Martín-Caro, 2024).

Los factores académicos y socioeconómicos fueron los más recurrentes en los estudios examinados, lo que confirmó el carácter multidimensional de la deserción universitaria. Las calificaciones, las asignaturas reprobadas y el progreso académico se relacionaron con condiciones económicas y familiares que afectaron la continuidad estudiantil. Fernández-Martín et al. (2019) identificaron que la probabilidad de abandono dependía de la interacción de diversas variables académicas y contextuales. De manera similar, Milione et al. (2023) determinaron que los datos obtenidos al inicio de la carrera no resultaban suficientes y que la incorporación del

rendimiento académico incrementaba la capacidad predictiva. Por ello, un sistema institucional basado en un número limitado de indicadores podría omitir trayectorias de riesgo complejas.

El predominio del aprendizaje automático, empleado como enfoque principal en el 80 % de las publicaciones, coincidió con otras revisiones de minería de datos educativa. Los árboles de decisión fueron las técnicas más frecuentes, seguidos por las redes neuronales y *random forest*. Agrusti et al. (2019) también identificaron una utilización recurrente de árboles de decisión, métodos probabilísticos, redes neuronales y regresión logística, mientras que Parra-Sánchez et al. (2023) señalaron que la mayoría de los estudios sobre deserción se orientaba a tareas de clasificación. La presencia de algoritmos interpretables resulta pertinente en el ámbito educativo, debido a que facilita la identificación de las variables asociadas con cada alerta y favorece la toma de decisiones institucionales fundamentadas (Urbina-Nájera et al., 2020).

Las métricas reportadas mostraron que ningún algoritmo presentó una superioridad constante en todos los contextos. Kuz y Morales (2023) informaron una exactitud superior al 99 % mediante XGBoost, mientras que Rivera (2021) obtuvo valores cercanos al 88 % con KNN y árboles de decisión. Por su parte, Pérez-Gutiérrez (2020) destacó el desempeño de *random forest* en distintos momentos de la trayectoria académica. Estas diferencias deben interpretarse con cautela, debido a que los estudios utilizaron poblaciones, variables, tamaños muestrales y procedimientos de validación distintos. En consecuencia, la selección de un modelo no debería sustentarse únicamente en la exactitud informada, sino también en su capacidad de generalización, interpretabilidad y utilidad para orientar intervenciones concretas.

Los resultados institucionales evidenciaron una brecha entre las capacidades predictivas descritas en la literatura y el seguimiento aplicado en la UTEQ. Aunque la mayoría de los tutores utilizaba recursos digitales, el proceso dependía principalmente de herramientas ofimáticas, encuestas socioeconómicas y revisión manual de calificaciones. En contraste, los estudios revisados integraron variables académicas, económicas, personales y demográficas para clasificar niveles de riesgo. Hinojosa et al. (2022) indicaron que la agrupación de estudiantes según su probabilidad de abandono podía facilitar la organización de acciones de acompañamiento, mientras que Villarreal-Torres et al. (2024) demostraron la aplicabilidad de modelos supervisados para reconocer anticipadamente posibles casos de deserción.

La incorporación de inteligencia artificial en la UTEQ requeriría depurar y actualizar previamente los datos, integrar las fuentes institucionales y establecer protocolos que transformen las alertas en acciones de apoyo. La predicción no reduciría por sí sola la deserción, pues su utilidad dependería de tutorías, orientación académica, apoyo económico y seguimiento personalizado. Valero et al. (2022) señalaron que las variables académicas y socioeconómicas podían emplearse conjuntamente para anticipar el abandono, mientras que Jimbo-Santana et al. (2023) destacaron la

importancia de integrar estas tecnologías en la gestión educativa. No obstante, cualquier modelo debería validarse con información propia de la UTEQ y funcionar como apoyo al criterio profesional, sin sustituir la intervención humana (Alquinga-Collaguazo et al., 2024).

5. Conclusiones

La revisión evidenció que la inteligencia artificial posee un potencial relevante para fortalecer la predicción temprana de la deserción estudiantil en educación superior, especialmente mediante algoritmos de aprendizaje automático. Los factores académicos y socioeconómicos fueron los más recurrentes, mientras que los árboles de decisión, las redes neuronales y *random forest* concentraron la mayor aplicación. En la UTEQ, el seguimiento académico continuó sustentándose principalmente en herramientas ofimáticas, encuestas socioeconómicas y revisión de calificaciones, sin sistemas automatizados de alerta. Esta brecha mostró la necesidad de depurar, integrar y actualizar los datos institucionales antes de implementar modelos predictivos. Su incorporación debería complementar la labor tutorial, facilitar intervenciones oportunas y validarse previamente con información propia, bajo criterios técnicos, éticos y de protección de datos.

CONFLICTO DE INTERESES

“Los autores declaran no tener ningún conflicto de intereses”.

Referencias Bibliográficas

- Agrusti, F., Bonavolontà, G., & Mezzini, M. (2019). University dropout prediction through educational data mining techniques: A systematic review. *Journal of e-Learning and Knowledge Society*, 15(3), 161–182. <https://doi.org/10.20368/1971-8829/1135017>
- Alquinga-Collaguazo, N. C., Andino-Jaramillo, R. A., & Guerrero-Bermúdez, Á. E. (2024). Percepciones, usos y desafíos pedagógicos del uso de la inteligencia artificial y rendimiento académico en estudiantes de educación superior. *Revista Científica Enfoques del Conocimiento*, 1(1), 1–18. <https://doi.org/10.55813/gaea/revistacec/v1/n1/7>
- Casanova-Villalba, C. I., Herrera-Sánchez, M. J., & Proaño-González, E. A. (2023). Impacto de la analítica predictiva en la toma de decisiones gerenciales. *Revista Científica Ciencia y Método*, 1(3), 16–30. <https://doi.org/10.55813/gaea/rcym/v1/n3/17>
- Cruz, E., González, M., & Rangel, J. C. (2022). Técnicas de machine learning aplicadas a la evaluación del rendimiento y a la predicción de la deserción de

- estudiantes universitarios, una revisión. *Prisma Tecnológico*, 13(1), 77–87. <https://doi.org/10.33412/pri.v13.1.3039>
- Eckert, K. B., & Suénaga, R. (2015). Análisis de deserción-permanencia de estudiantes universitarios utilizando técnica de clasificación en minería de datos. *Formación Universitaria*, 8(5), 3–12. <https://doi.org/10.4067/S0718-50062015000500002>
- Fernández-Martín, T., Solís-Salazar, M., Hernández-Jiménez, M. T., & Moreira-Mora, T. E. (2019). Un análisis multinomial y predictivo de los factores asociados a la deserción universitaria. *Revista Electrónica Educare*, 23(1), 1–25. <https://doi.org/10.15359/ree.23-1.5>
- Hinojosa, M., Derpich, I., Alfaro, M., Ruete, D., Caroca, A., & Gatica, G. (2022). Procedimiento de agrupación de estudiantes según riesgo de abandono para mejorar la gestión estudiantil en educación superior. *Texto Livre*, 15, e37275. <https://doi.org/10.35699/1983-3652.2022.37275>
- Jimbo-Santana, P., Lanzarini, L. C., Jimbo-Santana, M., & Morales-Morales, M. (2023). Inteligencia artificial para analizar el rendimiento académico en instituciones de educación superior. Una revisión sistemática de la literatura. *Cátedra*, 6(2), 30–50. <https://doi.org/10.29166/catedra.v6i2.4408>
- Kuz, A., & Morales, R. (2023). Ciencia de datos educativos y aprendizaje automático: un caso de estudio sobre la deserción estudiantil universitaria en México. *Education in the Knowledge Society (EKS)*, 24, e30080. <https://doi.org/10.14201/eks.30080>
- Martín-Caro, E. (2024). Análisis de estrategias innovadoras para retención estudiantil con inteligencia artificial: una perspectiva multidisciplinaria. *European Public & Social Innovation Review*, 9, 1–20. <https://doi.org/10.31637/epsir-2024-440>
- Milione, H. F., Fernández, D., Ortiz, C. A., Giménez Prieto, B., Magariños, M. S., & Appeceix, J. S. (2023). Modelo predictivo para estimar la deserción de estudiantes en la carrera de medicina de la Universidad Nacional de La Matanza. *Revista de Investigaciones del Departamento de Ciencias de la Salud*, 2(1), 33–40. <https://doi.org/10.54789/rs.v2i1.20>
- Miranda, M. A., & Guzmán, J. (2017). Análisis de la deserción de estudiantes universitarios usando técnicas de minería de datos. *Formación Universitaria*, 10(3), 61–68. <https://doi.org/10.4067/S0718-50062017000300007>
- Okuda Benavides, M., & Gómez-Restrepo, C. (2005). Métodos en investigación cualitativa: triangulación. *Revista Colombiana de Psiquiatría*, 34(1), 118–124. https://www.scielo.org.co/scielo.php?script=sci_arttext&pid=S0034-74502005000100008
- Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., ... Alonso-Fernández, S. (2021). Declaración PRISMA 2020: una guía actualizada para la publicación de revisiones sistemáticas. *Revista Española de Cardiología*, 74(9), 790–799. <https://doi.org/10.1016/j.recesp.2021.06.016>

- Parra-Sánchez, J. S., Torres Pardo, I. D., & Martínez de Merino, C. Y. (2023). Factores explicativos de la deserción universitaria abordados mediante inteligencia artificial. *Revista Electrónica de Investigación Educativa*, 25, e18, 1–17. <https://doi.org/10.24320/redie.2023.25.e18.4455>
- Pérez-Gutiérrez, B. R. (2020). Comparación de técnicas de minería de datos para identificar indicios de deserción estudiantil, a partir del desempeño académico. *Revista UIS Ingenierías*, 19(1), 193–204. <https://doi.org/10.18273/revuin.v19n1-2020018>
- Piedra-Castro, W. I., Cajamarca-Correa, M. A., Burbano-Buñay, E. S., & Moreira-Alcívar, E. F. (2024). Integración de la inteligencia artificial en la enseñanza de las ciencias sociales en la educación superior. *Journal of Economic and Social Science Research*, 4(3), 105–126. <https://doi.org/10.55813/gaeal/jessr/v4/n3/123>
- Rivera Vergaray, K. (2021). Modelo predictivo para la detección temprana de estudiantes con alto riesgo de deserción académica. *Innovación y Software*, 2(2), 6–13. <https://doi.org/10.48168/innosoft.s6.a40>
- Sanhueza Gutiérrez, D., King-Domínguez, A., & Améstica-Rivas, L. (2021). Incidencia de la gestión universitaria en la deserción estudiantil de las universidades públicas en Chile. *IE Revista de Investigación Educativa de la REDIECH*, 12, e1270. https://doi.org/10.33010/ie_rie_rediech.v12i0.1270
- Santos, C. M. da C., Pimenta, C. A. de M., & Nobre, M. R. C. (2007). The PICO strategy for the research question construction and evidence search. *Revista Latino-Americana de Enfermagem*, 15(3), 508–511. <https://doi.org/10.1590/S0104-11692007000300023>
- Secretaría Nacional de Planificación. (2021). *Plan de Creación de Oportunidades 2021–2025*. <https://www.planificacion.gob.ec/plan-de-creacion-de-oportunidades-2021-2025/>
- Urbina-Nájera, A. B., Camino-Hampshire, J. C., & Cruz Barbosa, R. (2020). Deserción escolar universitaria: patrones para prevenirla aplicando minería de datos educativa. *RELIEVE. Revista Electrónica de Investigación y Evaluación Educativa*, 26(1), 1–19. <https://doi.org/10.7203/relieve.26.1.16061>
- Valero Cajahuanca, J. E., Navarro Raymundo, Á. F., Larios Franco, A. C., & Julca Flores, J. D. (2022). Deserción universitaria: evaluación de diferentes algoritmos de machine learning para su predicción. *Revista de Ciencias Sociales*, 28(3), 362–375. <https://doi.org/10.31876/rcs.v28i3.38480>
- Villarreal-Torres, H., Ángeles-Morales, J., Marín-Rodríguez, W., & Cano-Mejía, J. (2024). Modelo de clasificación para la deserción estudiantil en las universidades públicas del Perú. *Revista de Ciencias Sociales*, 30(1), 452–469. <https://doi.org/10.31876/rcs.v30i1.41667>