ISBN: 978-628-96613-3-0. ISSN: 2414-6390. Digital Object Identifier: https://dx.doi.org/10.18687/LACCEI2026.1.1.1322

Quantitative prediction of the risk of failure using learning analytics in virtual engineering courses

Henri Emmanuel Lopez Gomez[1]; Roberto Carlos Davila Moran[2]; Vilma Luz Aparicio Salas[3]; Zoraida Loaiza Ortiz[4]; Juan Manuel Sanchez Soto[5]; Julia Marleni Martin Marcelo[6]; Dimna Zoila Alfaro Quezada[7]
[1]Universidad Tecnológica del Perú, Perú, *c31648@utp.edu.pe*
[2]Universidad Continental, Perú, *rdavilam@continental.edu.pe*
[3,4]Universidad Nacional San Antonio Abad del Cusco, Perú, *vilma.aparicio@unsaac.edu.pe*, *zoraida.loaiza@unsaac.edu.pe*
[5,6]Universidad Peruana Los Andes, Perú, *d.jsanchezs@ms.upla.edu.pe*, *d.jmartin@ms.upla.edu.pe*
[7]Universidad Privada San Juan Bautista, Perú, *dimna.alfaro@upsjb.edu.pe*

Abstract– ***The expansion of online education in engineering programs has been accompanied by persistently high failure rates, making data-driven early warning systems urgently needed. This article presents and evaluates a quantitative, interpretable, and reproducible methodological framework for predicting the risk of failing introductory engineering courses delivered online, based on learning analytics obtained from the Learning Management System (LMS) and the academic system. The approach was applied to a group of 200 students and included variables such as platform activity (number of active days, logins, resource views, and submissions); performance at a later stage (cumulative GPA and percentage of assessments submitted); and prior academic performance (GPA and number of course re-enrollments). Supervised logistic regression, random forest, and gradient reinforcement models were compared and evaluated based on AUC, sensitivity, specificity, F1 score, and Brier score. The probabilities of failing were transformed into a risk score (low/medium/high). The best-performing model showed an AUC of 0.84 at week 5 for the logistic regression model and approached 0.90 for gradient boosting with very good calibration. Risk categorization was able to accurately group approximately 60% of all students who ultimately failed into the high-risk category. (This high-risk category represented only about 25% of the cohort.) These results demonstrate the potential operational capability of the early warning system. The framework provides an interpretable tool for faculty and administrators to categorize risk levels in an effort to improve early intervention practices based on student performance data, as well as to enhance data-driven decision-making in the field of virtual engineering education.***

Keywords-- ***learning analytics, risk prediction, early warning systems, engineering education, virtual courses.***

Predicción cuantitativa del riesgo de desaprobación mediante analíticas de aprendizaje en cursos de ingeniería en modalidad virtual

Henri Emmanuel Lopez Gomez[1]; Roberto Carlos Davila Moran[2]; Vilma Luz Aparicio Salas[3]; Zoraida Loaiza Ortiz[4]; Juan Manuel Sanchez Soto[5]; Julia Marleni Martin Marcelo[6]; Dimna Zoila Alfaro Quezada[7]
[1]Universidad Tecnológica del Perú, Perú, *c31648@utp.edu.pe*
[2]Universidad Continental, Perú, *rdavilam@continental.edu.pe*
[3,4]Universidad Nacional San Antonio Abad del Cusco, Perú, *vilma.aparicio@unsaac.edu.pe*, *zoraida.loaiza@unsaac.edu.pe*
[5,6]Universidad Peruana Los Andes, Perú, *d.jsanchezs@ms.upla.edu.pe*, *d.jmartin@ms.upla.edu.pe*
[7]Universidad Privada San Juan Bautista, Perú, *dimna.alfaro@upsjb.edu.pe*

Resumen– ***La expansión de la educación en línea en programas de ingeniería ha venido acompañada de tasas de desaprobación persistentemente altas, lo que hace urgente contar con sistemas de alerta temprana basados en datos. Este artículo presenta y evalúa un marco metodológico cuantitativo, interpretable y reproducible para predecir el riesgo de desaprobación en cursos básicos de ingeniería impartidos en modalidad virtual, a partir de analíticas de aprendizaje obtenidas del LMS y del sistema académico. El enfoque se aplicó en un grupo de 200 estudiantes e incluyó las variables de actividad en la plataforma (número de días activos, inicios de sesión, visualizaciones de recursos y envíos); rendimiento en una etapa posterior (GPA acumulativo y porcentaje de evaluaciones enviadas); y rendimiento académico previo (GPA y número de reinscripciones en cursos). Se compararon y evaluaron modelos supervisados de regresión logística, bosque aleatorio y refuerzo de gradiente basándose en el AUC, la sensibilidad, la especificidad, la puntuación F1 y la puntuación Brier. Las probabilidades de desaprobación se transformaron en una puntuación de riesgo (baja/media/alta). El modelo con mejor rendimiento mostró un AUC de 0,84 en la semana 5 para el modelo de regresión logística y se acercó a 0,90 para el refuerzo de gradiente con una calibración muy buena. La categorización del riesgo fue capaz de agrupar con precisión aproximadamente al 60 % de todos los estudiantes que finalmente desaprobaron en la categoría de alto riesgo. (Esta categoría de alto riesgo solo representaba aproximadamente el 25 % de la cohorte). Estos resultados demuestran la capacidad operativa potencial del sistema de alerta. El marco proporciona una herramienta interpretable para que los profesores y administradores cataloguen los niveles de riesgo en un esfuerzo por mejorar las prácticas de intervención temprana basadas en los datos de rendimiento de los estudiantes, así como para mejorar la toma de decisiones basada en datos en el ámbito de la educación virtual en ingeniería.***

Palabras clave-- analíticas de aprendizaje, predicción de riesgo, sistemas de alerta temprana, educación en ingeniería, cursos virtuales.

I. Introducción

En la última década, la educación superior ha experimentado un crecimiento sostenido de la oferta de programas híbridos y completamente virtuales, incluidos los programas de ingeniería. Al mismo tiempo, la literatura documenta que las tasas de deserción y desaprobación en educación a distancia continúan siendo elevadas y heterogéneas entre países e instituciones [1], [2].

Paralelamente, el campo de learning analytics (LA) ha evolucionado como una línea de investigación aplicada que integra la minería de datos educativos, la estadística y el machine learning para tratar de entender y mejorar los procesos de enseñanza-aprendizaje. Las revisiones sistemáticas más recientes muestran un aumento evidente en la aplicación de modelos de machine learning y, más recientemente, sobre todo, de inteligencia artificial generativa para respaldar al proceso de decisiones en la educación superior, especialmente en la predicción de rendimiento, la predicción de abandono y el éxito académico [3].

En este contexto, los datos generados por los entornos virtuales de aprendizaje (p. ej., Moodle, Canvas, Brightspace) ofrecen un registro detallado de la actividad de los estudiantes: accesos, tiempo de permanencia, visualización de recursos, participación en foros y entregas de evaluaciones. Diversos estudios han mostrado que estos registros, combinados con información académica previa y datos demográficos básicos, permiten desarrollar modelos predictivos capaces de identificar estudiantes en riesgo con niveles de desempeño que suelen superar los de modelos basados únicamente en calificaciones finales o promedios acumulados [2], [4], [5].

En los estudios de ingeniería, las asignaturas que se consideran fundamentales (Cálculo, Física o Mecánica de Fluidos) poseen una elevada tasa histórica de desaprobación y repetición, lo que sugiere que son candidatos ideales para aplicar herramientas de learning analytics. En esta línea, el estudio más reciente realizado en la universidad, en el marco de la asignatura Mecánica de Fluidos, analizó casi 68000 registros de interacción en el campus virtual y mostró que tanto la intensidad de uso de esta plataforma como la evolución de la actividad a lo largo del semestre encuentran una estrecha vinculación con el éxito o el fracaso en la materia [6].

A pesar de que la evidencia reciente avala la idoneidad de los modelos predictivos orientados a la detección temprana del abandono y la desaprobación basados en datos de LMS y registros institucionales, al menos tres limitaciones relevantes para el contexto de la ingeniería bajo la modalidad virtual

pueden identificarse. En primer lugar, se observa un enfoque institucional y poca granularidad a nivel de curso: los anteriores estudios se sustentan en datos provenientes de múltiples programas o facultades y modelan el riesgo de abandono a nivel institucional y con horizontes temporales de múltiples semestres [1], [2]. Si bien estos enfoques son valiosos para la gestión estratégica, ofrecen menos detalles sobre los patrones específicos de interacción y evaluación en cursos concretos. En segundo lugar, existe un predominio de modelos de "caja negra"; investigaciones recientes reportan el uso exitoso de algoritmos de *gradient boosting*, redes neuronales profundas y modelos de ensamblado para la predicción de abandono, con métricas de desempeño destacadas [5], [7], [8]. Sin embargo, la opacidad de estos modelos dificulta que docentes y responsables académicos comprendan por qué un estudiante es clasificado como "en riesgo" y cómo actuar pedagógicamente sobre esa información [8], [9]. Finalmente, se aprecia una escasa especificidad para cursos de ingeniería completamente virtuales. Aunque existen estudios de caso en ingeniería que aprovechan analíticas de aprendizaje para caracterizar la interacción con el campus virtual y las actividades de evaluación, pocos trabajos desarrollan marcos de alerta temprana cuantitativos y explícitamente replicables para cursos básicos de ingeniería en modalidad totalmente en línea [6], [10].

En respuesta a estas brechas, el presente artículo tiene como objetivo general proponer un marco metodológico cuantitativo, interpretable y reproducible para la predicción temprana del riesgo de desaprobación en cursos de ingeniería impartidos en modalidad virtual, basado en analíticas de aprendizaje. En coherencia con este objetivo, se abordan tres preguntas de investigación. La primera (RQ1) se orienta a identificar qué variables derivadas de los registros de un LMS y del historial académico describen con mayor fuerza el riesgo de desaprobar un curso virtual de ingeniería. La segunda (RQ2) busca determinar qué nivel de desempeño discriminante puede alcanzarse aprovechando los modelos supervisados clásicos (como la regresión logística y los bosques aleatorios) para predecir la desaprobación a partir de datos recolectados en las primeras semanas del curso. La tercera (RQ3) se centra en cómo se puede traducir la probabilidad de riesgo estimada en indicadores categóricos comprensibles y aplicables para el profesorado y el equipo de soporte académico.

De estas preguntas se derivan tres contribuciones principales. Primero, se propone un pipeline cuantitativo para extraer, transformar y analizar datos de cursos virtuales de ingeniería, integrando buenas prácticas de anonimización y gobernanza de datos del alumnado. Segundo, se presenta un esquema de modelado del riesgo que combina transparencia analítica y capacidad predictiva**,** complementando la regresión logística con bosques aleatorios y gradient boosting. Tercero, se diseña un esquema de categorización de riesgo (bajo, medio, alto) vinculado a métricas estándar de validación (AUC, sensibilidad, especificidad, F1 y calibración), orientado a facilitar su uso pedagógico y su adopción por parte de docentes y responsables académicos.

II. MARCO TEÓRICO

A. Learning analytics y sistemas de alerta temprana

El campo de *learning analytics* se centra en la medición, recopilación, análisis y reporte de datos sobre los estudiantes y sus contextos, con el propósito de comprender y optimizar el aprendizaje y los entornos en los que ocurre [3], [8]. Una línea central dentro de LA y de la minería de datos educativos (EDM) es el desarrollo de sistemas de alerta temprana (*Early Warning Systems*, EWS) que permitan identificar estudiantes en riesgo con suficiente anticipación para implementar intervenciones efectivas.

Desde una perspectiva conceptual, los sistemas de alerta temprana en educación superior pueden entenderse al menos desde tres enfoques complementarios. El primero se orienta a la predicción del riesgo como problema de clasificación, priorizando métricas de desempeño y capacidad discriminante. El segundo enfatiza la interpretabilidad y explicabilidad de los modelos, reconociendo que la utilidad pedagógica de una alerta depende de que los actores educativos comprendan qué señales la originan y cómo pueden intervenir sobre ellas. El tercero pone el foco en la traducción institucional de las analíticas, es decir, en cómo los resultados del modelo se integran en dashboards, procesos de tutoría, asesoría académica y toma de decisiones docentes. Esta distinción resulta importante porque un sistema puede ser estadísticamente preciso, pero pedagógicamente poco útil si no ofrece indicadores comprensibles y accionables para el acompañamiento del estudiantado.

Russell et al. [11] describen *Elements of Success*, una plataforma analítica de aprendizaje que proporciona a estudiantes etiquetados como "en riesgo" retroalimentación periódica sobre su desempeño y una estimación actualizada de su calificación, lo que, junto con otros estudios similares [12], [13], demuestra el potencial de los EWS para apoyar la resiliencia académica en cursos introductorios de ciencias.

En el ámbito de la ingeniería y áreas afines, estudios como los de West et al. [10] y Ngulube y Ncube [14] combinaron indicadores de LA con percepciones estudiantiles para explorar los patrones de interacción en un curso en línea de gestión de la construcción, mostrando que las métricas de interacción con el contenido y de autopercepción de aprendizaje se relacionan de manera significativa con el desempeño final.

Un patrón consistente en la literatura es que la detección temprana del riesgo, idealmente en las primeras semanas del curso, incrementa la probabilidad de que las intervenciones (tutorías, acompañamiento, rediseño de actividades) resulten efectivas [4], [8]. De ahí la importancia de marcos predictivos que no solo sean precisos, sino también oportunos y fácilmente interpretables por su audiencia principal: docentes y coordinadores de programa.

B. Modelos predictivos de abandono y desaprobación

Los modelos predictivos aplicados a abandono y fracaso en educación superior utilizan, en general, tres tipos de variables. En primer lugar, se consideran las características de entrada, que incluyen datos de admisión, rendimiento previo y variables sociodemográficas. En segundo lugar, se incorporan los datos académicos de progreso, tales como calificaciones parciales, créditos aprobados e historial de repetición, que permiten capturar la trayectoria académica del estudiante a lo largo del tiempo [2], [8]. En tercer lugar, se emplean las interacciones con el LMS, entre las que se encuentran la frecuencia de accesos, el tiempo de permanencia en la plataforma, la entrega de actividades, la participación en foros y la visualización de recursos, las cuales proporcionan una aproximación detallada al comportamiento de estudio en entornos virtuales [4], [5], [6].

Seo et al. [1] desarrollan un modelo de predicción de abandono en una universidad a distancia en Corea del Sur, integrando estas tres fuentes de datos. Usando un enfoque de selección de variables por eliminación hacia atrás, identifican como indicadores clave de riesgo variables como edad, área de residencia, ocupación, GPA y métricas de actividad en el LMS; además, muestran que los algoritmos de *gradient boosting* logran el mejor desempeño, mientras que la regresión logística ofrece interpretaciones más transparentes con un rendimiento competitivo.

En un estudio reciente basado en registros de Moodle, Marcolino et al. [5] optimizan diferentes modelos de *machine learning* para la predicción de abandono, demostrando que la combinación de técnicas de *oversampling* para tratar el desbalance de clases y búsqueda exhaustiva de hiperparámetros produce modelos con alta capacidad predictiva en cursos técnicos.

Osborne y Lang [4], por su parte, se enfocan en la identificación de estudiantes en riesgo de reprobar cursos específicos a partir de datos del libro de calificaciones de un LMS durante las primeras cinco semanas de clases. Mediante una comparación de modelos (regresión logística, *k*-vecinos más cercanos, bosques aleatorios y una red neuronal simple), ofrecen evidencia de que es posible realizar una predicción temprana del riesgo de fracaso en un curso utilizando exclusivamente información de calificaciones parciales y de tareas no registradas.

Finalmente, investigaciones como las de Lee et al. [9] y Čotić Poturić et al. [8] comentan sobre la interpretabilidad de los modelos predictivos en la educación. Lee et al. [9] proponen un modelo Markoviano interpretable para analíticas predictivas en torno a la educación en línea y Čotić Poturić et al. [8] desarrollan un algoritmo de *scoring* para cursos STEM que transforma predictores en variables dicotómicas y utilizan medidas como V de Cramér y el índice de Youden para definir umbrales de riesgo interpretables y comunicables.

Estos trabajos permiten identificar una tensión recurrente en la literatura: mientras los modelos más complejos suelen alcanzar mejores niveles de ajuste y capacidad predictiva, los enfoques más interpretables resultan más útiles para contextos educativos donde la explicación de la alerta y la posibilidad de actuar pedagógicamente sobre ella son tan importantes como la métrica de desempeño. En consecuencia, la discusión actual no se limita a determinar qué algoritmo predice mejor, sino a establecer bajo qué condiciones un modelo puede considerarse suficientemente preciso, explicable y operativamente útil para apoyar decisiones en tiempo oportuno. En cursos de ingeniería impartidos completamente en línea, esta tensión adquiere especial relevancia, ya que la interacción del estudiantado con el LMS genera un volumen considerable de datos, pero no toda señal estadísticamente significativa se traduce automáticamente en una intervención pedagógica viable. En este marco, el presente estudio se sitúa en una posición intermedia: propone un enfoque predictivo orientado al riesgo académico que busca equilibrar capacidad discriminante, interpretabilidad y aplicabilidad práctica a nivel de asignatura.

A partir de esta revisión, la brecha que aborda el presente trabajo puede delimitarse con mayor precisión: todavía son limitados los estudios que, en cursos básicos de ingeniería impartidos completamente en modalidad virtual, combinen datos del LMS y antecedentes académicos para construir un sistema de alerta temprana que no solo alcance un buen desempeño predictivo, sino que además traduzca sus resultados en categorías de riesgo comprensibles para la docencia y la gestión académica. En consecuencia, el aporte del estudio no radica únicamente en comparar algoritmos, sino en proponer un marco metodológico que articule analíticas de aprendizaje, predicción temprana e interpretación pedagógica del riesgo académico.

III. Metodología

A. Diseño del estudio

Esta investigación se concibe como un estudio cuantitativo, explicativo y longitudinal, orientado a predecir, de forma binaria (aprobado/desaprobado), el desenlace final de una asignatura de ingeniería impartida completamente en formato virtual (es decir, alojada y gestionada en un LMS institucional). El curso combina evaluación continua (tareas, cuestionarios, proyectos) con al menos una evaluación sumativa que aporta una proporción relevante de la nota final. El tamaño de cada cohorte semestral oscila entre 80 y 250 alumnos, un rango habitual en cursos básicos de ingeniería [2], [6]. Aunque el marco se presenta de forma general, puede adaptarse a contextos académicos distintos (por ejemplo, Cálculo I, Programación o Circuitos Eléctricos) ajustando el conjunto de variables y/o la ventana temporal de predicción.

Este diseño resulta adecuado porque la predicción temprana del riesgo de desaprobación exige observar la evolución de indicadores de actividad y desempeño a lo largo del tiempo, en lugar de limitarse a una medición única al cierre del curso. La orientación cuantitativa permite estimar relaciones entre variables observables del LMS, antecedentes académicos y desenlace final, mientras que el carácter explicativo facilita identificar qué factores muestran mayor asociación con la desaprobación y con qué magnitud. Asimismo, el enfoque longitudinal permite construir ventanas

de predicción temprana coherentes con la lógica de los sistemas de alerta, donde la utilidad del modelo depende no solo de su precisión, sino también del momento en que la señal de riesgo puede detectarse y traducirse en acciones de apoyo. No obstante, este diseño no permite establecer causalidad estricta ni garantizar que el mismo desempeño predictivo se reproduzca automáticamente en otros cursos, cohortes o instituciones, por lo que sus resultados deben interpretarse en términos de capacidad predictiva contextualizada y no de generalización universal.

B. Fuentes de datos y variables

Se integran tres fuentes principales de datos, todas de naturaleza cuantitativa y recolectadas automáticamente, sin aplicación de cuestionarios ni instrumentos subjetivos. En primer lugar, se consideran los datos del LMS, que describen la actividad en el curso e incluyen el número total de accesos y de días activos, el tiempo total estimado en la plataforma (cuando el LMS lo permite), el número de recursos de contenido visualizados (PDF, videos, páginas), el número de actividades evaluativas vistas y los intentos enviados, la participación en foros (mensajes publicados, respuestas, lecturas) y distintos indicadores de regularidad, como rachas de días consecutivos con acceso y periodos de inactividad prolongada [4], [6], [8]. En segundo lugar, se incluyen datos de desempeño dentro del curso, tales como las calificaciones de cuestionarios y tareas por semana, el porcentaje de actividades evaluativas entregadas y la nota acumulada al cierre de semanas de referencia (típicamente semanas 3, 5 y 7) [4], [8]. En tercer lugar, se incorporan datos académicos de contexto, que abarcan el promedio ponderado acumulado (GPA) previo, el número de veces que el estudiante ha inscrito la asignatura, el número de créditos aprobados en semestres anteriores y cierta información básica de cohorte, como la distinción entre estudiantes de primer ingreso y de reingreso [2].

La variable objetivo se define como Y=1 cuando el estudiante desaprueba la asignatura (incluyendo calificación reprobatoria, retiro con nota de reprobación o ausencia al examen final cuando esta se considera desaprobación) y Y=0 cuando la aprueba. Esta definición es consistente con trabajos que tratan la desaprobación y el abandono del curso como parte de un mismo fenómeno de riesgo de desaprobación a nivel de asignatura, en tanto ambas situaciones expresan una trayectoria académica que culmina sin logro satisfactorio del curso [2], [4].

C. Preprocesamiento y ventana temporal

El *pipeline* propuesto comprende varias etapas encadenadas. En primer lugar, se realiza la extracción y anonimización de los datos del LMS y del sistema académico mediante procesos ETL automatizados, asignando identificadores anónimos a los estudiantes y excluyendo información sensible como nombres o direcciones IP [4], [6]. En segundo lugar, se definen ventanas temporales de predicción temprana (por ejemplo, semanas 3 y 5), siguiendo aproximaciones como la de Osborne y Lang, quienes muestran que a partir de la quinta semana las señales predictivas se vuelven más estables [4]. En tercer lugar, se lleva a cabo la ingeniería de variables, construyendo agregados por ventana tales como número de accesos acumulados, porcentaje de tareas entregadas o máxima inactividad en días consecutivos [5], [8]. Posteriormente, se aborda el tratamiento de datos faltantes, aplicando estrategias como imputación simple (mediana) o la creación de categorías explícitas "sin información", según el tipo de variable. Finalmente, cuando la proporción de desaprobados es muy baja o muy alta, se recurre al balanceo de clases mediante técnicas de remuestreo (como SMOTE) o ajuste de pesos de clase en los modelos, tal como recomiendan meta–análisis recientes sobre predicción de abandono con datos de LMS [2], [4], [5].

Con el fin de hacer explícito el pipeline analítico, el proceso metodológico se organizó en la siguiente secuencia: i) extracción y consolidación de registros del LMS y del sistema académico; ii) anonimización y asignación de identificadores codificados; iii) definición de ventanas temporales de predicción; iv) construcción de variables agregadas por ventana; v) tratamiento de datos faltantes e inconsistencias; vi) balanceo de clases cuando la distribución aprobado/desaprobado lo requirió; vii) entrenamiento y validación de modelos supervisados; y viii) traducción de probabilidades estimadas en categorías operativas de riesgo. Esta explicitación busca reforzar la reproducibilidad del marco propuesto y facilitar su adaptación en otros cursos virtuales de ingeniería con estructuras de datos comparables.

D. Modelos y métricas de evaluación

Para equilibrar precisión e interpretabilidad, el marco propone un conjunto básico de modelos supervisados. La regresión logística binaria se utiliza como modelo principal por su transparencia y por la facilidad que ofrece para interpretar los coeficientes en términos de *odds ratios*; se ajusta para cada ventana temporal de predicción e incorpora regularización (ridge o lasso) cuando el número de predictores es elevado [4], [8]. Como modelo complementario se emplea el bosque aleatorio, capaz de capturar interacciones no lineales entre predictores y de proporcionar medidas de importancia de variables —por impureza o por permutación— útiles para identificar los factores con mayor contribución al riesgo [4], [5], [8]. De manera opcional, se considera un modelo de gradient boosting, incluido como referencia cuando se dispone de recursos computacionales suficientes, dado su buen desempeño reportado en múltiples estudios de predicción de abandono [5].

El desempeño de los modelos se evalúa mediante validación cruzada estratificada y un conjunto de métricas estándar: área bajo la curva ROC (AUC), sensibilidad (recall de la clase "desaprueba"), especificidad, F1 y medidas de calibración, como las curvas de calibración y el Brier score, especialmente relevantes para interpretar probabilidades de riesgo en educación [3], [8]. Para las métricas dependientes de un umbral (sensibilidad, especificidad y F1), las probabilidades estimadas se binarizan usando un punto de corte fijo aplicado de forma consistente durante la evaluacion.

E. Traducción de probabilidades a categorías de riesgo

Siguiendo la lógica de los algoritmos de *scoring* en STEM, se propone convertir la probabilidad estimada de desaprobación en una puntuación de riesgo y, posteriormente, en categorías cualitativas [8], [11]. En un esquema de implementación, las probabilidades se normalizan en una escala de 0 a 100 y se emplea el análisis ROC para definir umbrales operativos. En particular, se estima un umbral p_2 para la categoría de alto riesgo maximizando el índice de Youden (sensibilidad + especificidad -1) al discriminar "desaprueba" vs. "aprueba". Luego, se define un umbral inferior p_1 (< p_2) para identificar un grupo de bajo riesgo, de forma que el intervalo intermedio represente una zona de riesgo medio o indeterminado. A partir de estos puntos de corte, se definen tres niveles de riesgo: bajo (score < p_1), medio ($p_1 \leq$ score < p_2) y alto (score $\geq p_2$). La selección definitiva de los umbrales se realiza en diálogo con los equipos docentes, equilibrando el costo de los falsos positivos —intervenciones innecesarias— y de los falsos negativos —estudiantes en riesgo no detectados—.

F. Consideraciones éticas y de gobernanza de datos

El marco establece que la institución debe informar explícitamente a los estudiantes que sus datos se usaran con fines de mejora educativa; describir con claridad como se procede a la anonimización y al almacenamiento de la información, y como estas prácticas se alinean con sus políticas internas y con la normativa de protección de datos aplicable [3], [6] y asegurar que las salidas del modelo (etiquetas de riesgo) se utilicen únicamente como insumo para el soporte y el acompañamiento académico, no para la imposición de sanciones.

IV. RESULTADOS

A. Descripción de la muestra

La muestra estuvo constituida por 200 estudiantes que estaban matriculados en la materia durante un semestre académico. De estos, 130 aprobaron la materia (65 %) y 70 la desaprobaron (35 %), corroborando así el carácter crítico de la asignatura como generadora de riesgo académico. La Tabla I recoge las principales características de la cohorte.

Por otro lado, el grupo tenía en promedio 21,4 años (DE = 3,2) y las mujeres representaban un 38 %. El promedio ponderado acumulado (GPA) previo fue de 13,1 puntos sobre una escala de 0 a 20. Esta media mostró diferencias marcadas entre los aprobados (13,8) y los desaprobados (12,0). Adicionalmente, el grupo que desaprobó se había matriculado en la asignatura en un mayor número de ocasiones (1,7 frente a 1,1 para los aprobados), lo que refuerza la idea de que en esta materia suele haber repetidores y se generan cuellos de botella en el plan de estudios.

TABLA I
Características generales de la muestra

Variable	Muestra total (n = 200)	Aprobados (n = 130)	Desaprobados (n = 70)
Edad (años), media ± DE	21,4 ± 3,2	21,0 ± 3,0	22,1 ± 3,5
Mujeres (%)	38%	40%	34%
GPA previo (0–20), media ± DE	13,1 ± 1,7	13,8 ± 1,4	12,0 ± 1,8
Veces inscrito el curso, media ± DE	1,3 ± 0,7	1,1 ± 0,4	1,7 ± 0,9

B. Patrones de actividad en el LMS y desempeño parcial

En relación con la actividad en el LMS hasta la semana 5, la Tabla II muestra diferencias claras entre los grupos de aprobados y desaprobados. Los estudiantes que aprobaron registraron, en promedio, 24,3 días activos en la plataforma frente a 15,6 días en el grupo reprobado, así como un número significativamente mayor de accesos totales (145,2 frente a 82,7). El porcentaje de tareas entregadas hasta la semana 5 fue del 88 % en el grupo aprobado y del 54 % en el grupo desaprobado, mientras que la nota acumulada parcial se situó en 71,5 % y 41,2 %, respectivamente.

Estas diferencias resultaron estadísticamente significativas en todos los casos ($p < 0,001$), lo que sugiere que la combinación de regularidad de acceso, cumplimiento de actividades y desempeño parcial proporciona una señal temprana robusta sobre el riesgo de desaprobación.

TABLA II
Actividad en el LMS y desempeño parcial hasta la semana 5

Variable	Aprobados media (DE)	Desaprobados media (DE)	p-valor
Días activos en el LMS	24,3 (6,8)	15,6 (7,3)	< 0,001
Accesos totales al curso	145,2 (60,4)	82,7 (51,9)	< 0,001
Tareas/cuestionarios entregados (%)	88,0 (12,5)	54,3 (20,1)	< 0,001
Nota acumulada parcial (% escala 0–100)	71,5 (11,3)	41,2 (15,7)	< 0,001

C. Desempeño de los modelos predictivos

La Tabla III presenta el desempeño de los modelos supervisados considerados (regresión logística, bosque aleatorio y *gradient boosting*) para dos ventanas temporales de predicción: semana 3 y semana 5. Tal como era esperable, el rendimiento mejora cuando se dispone de más información (semana 5), pero incluso en la semana 3 se alcanzan valores de AUC cercanos o superiores a 0,80 en los modelos de ensamble.

En la semana 3, la regresión logística obtuvo un AUC de 0,78, con sensibilidad de 0,77 y especificidad de 0,71. El bosque aleatorio mejoró estos resultados con un AUC de 0,82, mientras que el modelo de *gradient boosting* alcanzó un AUC de 0,84 y un Brier score de 0,16. En la semana 5, la regresión logística elevó su AUC hasta 0,84; el bosque aleatorio alcanzó 0,88 y el *gradient boosting* llegó a 0,90, con sensibilidades entre 0,81 y 0,86 y Brier scores en el rango de 0,12 a 0,15. Como se observa en la Fig. 1, las curvas ROC de los tres modelos se sitúan claramente por encima de la diagonal aleatoria, destacando el modelo de gradient boosting, seguido del bosque aleatorio y de la regresión logística.

TABLA III
Desempeño de los modelos de predicción de desaprobación

Modelo	Ventana	AUC	Sensibilidad	Especificidad	F1	Brier
Regresión	Semana 3	0,78	0,77	0,71	0,69	0,19

logística						
Bosque aleatorio	Semana 3	0,82	0,79	0,76	0,73	0,17
Gradient boosting	Semana 3	0,84	0,80	0,78	0,75	0,16
Regresión logística	Semana 5	0,84	0,81	0,77	0,74	0,15
Bosque aleatorio	Semana 5	0,88	0,84	0,81	0,79	0,13
Gradient boosting	Semana 5	0,90	0,86	0,83	0,81	0,12

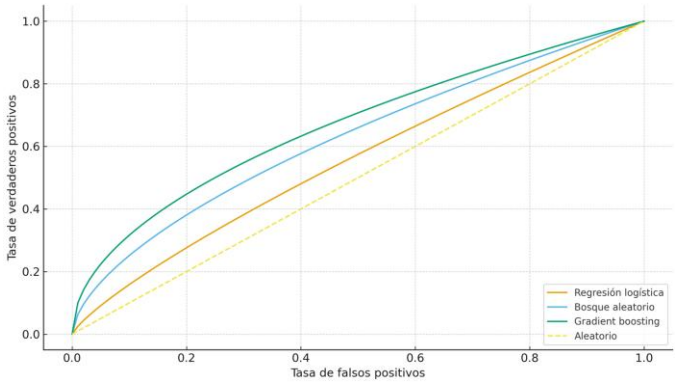

Fig. 1
Curvas ROC – ventana de predicción semana 5

D. Importancia de variables y perfil de riesgo

El análisis de importancia de variables en el bosque aleatorio sugiere que los predictores más influyentes son el porcentaje de tareas/cuestionarios entregados hasta la semana 5, la nota acumulada parcial, el número de días activos en el LMS y el GPA previo. Variables de interacción más finas, como los periodos de inactividad prolongada o el número de accesos en la semana inmediatamente anterior a una evaluación clave, también aportan información adicional, aunque en menor medida.

Los coeficientes de la regresión logística apuntan en la misma dirección: por cada incremento de 10 puntos porcentuales en la proporción de tareas entregadas, la razón de momios (odds) de desaprobar se reduce aproximadamente a la mitad, manteniendo constante el resto de variables. De forma análoga, un aumento de una desviación estándar en la nota parcial se asocia con una reducción sustantiva del riesgo de desaprobación. Esta convergencia entre modelos lineales e indicadores de importancia en modelos de ensamble refuerza la robustez de los hallazgos.

E. Categorización del riesgo y distribución de casos

A partir del modelo de regresión logística en la semana 5, la probabilidad de desaprobación se transformó en un *score* de riesgo entre 0 y 100 y se definieron tres categorías mediante puntos de corte derivados del análisis ROC y criterios operativos: riesgo bajo (score < 30), riesgo medio (30 ≤ score < 60) y riesgo alto (score ≥ 60).

Aplicando estos umbrales a la cohorte, el 25 % de los estudiantes (n = 50) se clasificó como de riesgo alto, el 30 % (n = 60) como de riesgo medio y el 45 % restante (n = 90) como de riesgo bajo. Sin embargo, la distribución del resultado final no fue homogénea entre categorías: el grupo de riesgo alto concentró 42 de los 70 desaprobados (60 %), el grupo de riesgo medio 22 casos (31 %) y el grupo de riesgo bajo solo 6 casos (9 %).

En términos prácticos, esto implica que intervenciones focalizadas sobre el 25 % de la cohorte etiquetada como de riesgo alto permitirían actuar tempranamente sobre alrededor del 60 % de los estudiantes que finalmente desaprueban el curso, con un volumen de falsos positivos manejable. La Fig. 2 resume gráficamente esta distribución, mostrando para cada categoría de riesgo el número de estudiantes aprobados y desaprobados.

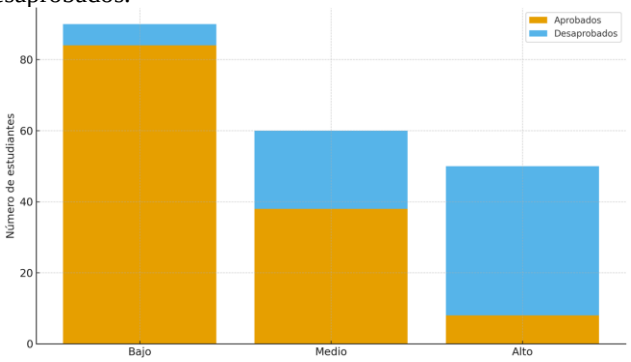

Fig. 2
Distribución de estudiantes por categoría de riesgo y resultado final

V. Discusión

A. Síntesis de los hallazgos principales

Este estudio mostró que los patrones tempranos de actividad en el LMS y el desempeño parcial se asocian de manera fuerte y consistente con el riesgo de desaprobación en un curso básico de ingeniería en modalidad virtual. Los estudiantes que desaprobaron presentaron menos días activos, menos accesos totales, menor proporción de tareas entregadas y notas parciales sensiblemente más bajas que las de sus pares que aprobaron. Estos resultados confirman que el comportamiento de estudio observable en las primeras semanas del curso constituye un indicador robusto del desenlace final.

En términos predictivos, los modelos supervisados alcanzaron niveles de desempeño comparables o superiores a los reportados en la literatura reciente: la regresión logística obtuvo AUC de 0,78 en la semana 3 y 0,84 en la semana 5, mientras que el *gradient boosting* llegó a 0,84 y 0,90 en las mismas ventanas. La categorización de la probabilidad de desaprobación en tres niveles (bajo, medio y alto) permitió concentrar alrededor del 60 % de los estudiantes finalmente desaprobados en el grupo de mayor riesgo, abarcando solo el 25 % de la cohorte. Esta combinación de poder discriminante y parsimonia sugiere que el marco propuesto es adecuado para sustentar decisiones de intervención temprana a nivel de asignatura.

B. Relación con la literatura sobre modelos predictivos y analíticas de aprendizaje

Los resultados obtenidos se alinean con estudios que han demostrado el potencial de los registros del LMS, combinados

con información académica previa, para predecir abandono y fracaso en educación superior. Marcolino et al. [5] mostraron que algoritmos de *machine learning* entrenados sobre datos de Moodle alcanzan métricas de precisión elevadas al modelar el riesgo de deserción, particularmente cuando se optimizan hiperparámetros y se atiende al desbalance de clases. De forma similar, Vaarma et al. [15] y Ohkawauchi et al. [16] evidencian que la interacción en el LMS, medida a través de accesos, entregas y tiempo en plataforma, es un predictor clave del riesgo de no aprobar cursos en línea.

El desempeño de los modelos en este trabajo es coherente con los AUC reportados por Seo et al. [1] en una universidad a distancia, quienes alcanzan valores entre 0,75 y 0,88 al integrar datos académicos y de actividad en línea en modelos de regresión y ensamble. Asimismo, el incremento del rendimiento cuando se amplía la ventana temporal de observación coincide con hallazgos de Goren et al. [17] quienes señalan que la disponibilidad de calificaciones parciales y comportamiento acumulado mejora sustancialmente la capacidad de detección temprana del riesgo.

En cuanto a la forma de operacionalizar el riesgo, el uso de un *score* continúo transformado en categorías discretas se inspira en propuestas recientes de algoritmos de *scoring* para cursos STEM. Čotić Poturić et al. [8] desarrollan un sistema que combina variables académicas y de comportamiento para clasificar a los estudiantes en niveles de riesgo, demostrando que la interpretación de los puntos de corte es tan importante como la métrica global de exactitud. El enfoque adoptado en este estudio se sitúa en esa misma línea, priorizando la coherencia con la capacidad real de respuesta de los docentes.

C. Interpretabilidad, explicabilidad y uso pedagógico de los modelos

Un aporte central de este trabajo radica en mostrar que la regresión logística, complementada con modelos de ensamble, ofrece un equilibrio adecuado entre precisión e interpretabilidad. La literatura reciente ha enfatizado la necesidad de modelos explicables en analíticas de aprendizaje, dado que los usuarios finales —docentes, asesores académicos y estudiantes— deben comprender las razones detrás de las alertas para poder actuar sobre ellas. Linden et al. [18], por ejemplo, utiliza analíticas explicables para identificar estudiantes desenganchados y subraya que la transparencia del modelo facilita la aceptación por parte del profesorado.

De igual forma, trabajos sobre tableros y plataformas de analíticas como *Elements of Success* de Russell et al. [19] o los dashboards de orientación académica de Gutiérrez et al. [20], evidencian que la utilidad percibida se incrementa si los indicadores son claros, accionables y relacionados con estrategias concretas de apoyo. En el presente estudio, la detección de las variables críticas—porcentaje de tareas entregadas, nota parcial, días activos y GPA previo— produce un mensaje explícito: la regularidad en participar y la implicación en la realización de las evaluaciones formativas durante las primeras semanas del curso son variables críticas para evitar la desaprobación.

Ese énfasis en la interpretación, a su vez, también se relaciona con las propuestas de marcos socio-técnicos sobre sistemas de alerta temprana, que advierten de los peligros de ceder las decisiones educativas a modelos de "caja negra" desconectados del contexto institucional y de las capacidades de intervención.

D. Implicaciones para la docencia y la gestión en ingeniería en modalidad virtual

Los hallazgos tienen implicaciones directas para la enseñanza y la gestión de cursos básicos de ingeniería impartidos en línea. En primer lugar, confirman que la ventana de las primeras cinco semanas constituye un periodo crítico para la identificación del riesgo. Esto coincide con revisiones recientes que señalan la importancia de sistemas de alerta capaces de actuar en las fases iniciales del semestre para maximizar el impacto de las intervenciones [3], [12], [13].

En segundo lugar, la clasificación de categorías de riesgo es una forma de poner en práctica esfuerzos de priorización: apoyar intensivamente —mediante tutorías personalizadas y un seguimiento estrecho— a los estudiantes del grupo de mayor riesgo; aportar un acompañamiento moderado —con recordatorios, recursos complementarios y mensajes personalizados — a quienes estarían en riesgo medio; y, finalmente, llevar a cabo una supervisión ligera, sustentada en un feedback automatizado, en el grupo de bajo riesgo. Esta idea de escalonamiento coincide con la recomendación de los trabajos más recientes basados en sistemas de alertas tempranas, un tipo de intervención que persigue la racionalización de las cargas de trabajo de docentes y asesores [8], [11], [18], [20].

En tercer lugar, los resultados muestran que hay que implementar los cursos de manera que recojan de forma sistemática evidencias de participación y aprendizaje. Estudios recientes sobre el análisis de datos en Moodle muestran, por ejemplo, que la utilidad de las analíticas depende de que las actividades y los recursos estén organizados para construir señales claras, comparables y útiles a lo largo del tiempo. Esto implica que las decisiones de diseño instruccional (como la frecuencia de las evaluaciones, el uso de foros o la introducción de cuestionarios) afectan a la calidad de las analíticas de aprendizaje, sin que estas acciones puedan considerarse neutras.

Desde un punto de vista operativo, la utilidad del sistema de alerta temprana depende de su traducción en acciones concretas de acompañamiento. En una implementación real, los estudiantes clasificados en riesgo alto podrían ser derivados a tutorías personalizadas, revisión prioritaria de entregas pendientes, contacto temprano por parte del docente o del equipo de soporte académico y seguimiento semanal de su participación en el LMS. En el caso del riesgo medio, las acciones podrían concentrarse en recordatorios focalizados, recursos de refuerzo, mensajes personalizados y monitoreo de cumplimiento en actividades clave. Para el grupo de riesgo bajo, el sistema permitiría mantener una supervisión ligera basada en retroalimentación automatizada y seguimiento general del progreso. Esta lógica de intervención escalonada

contribuye a que la predicción no se reduzca a una clasificación estadística, sino que funcione como insumo para priorizar decisiones pedagógicas y distribuir de manera más eficiente los recursos de apoyo académico.

E. Limitaciones y líneas de investigación futura

Este estudio presenta varias limitaciones que deben considerarse al interpretar los resultados. En primer lugar, los análisis se basan en un único curso de ingeniería en modalidad virtual; por tanto, el desempeño alcanzado por los modelos debe interpretarse como evidencia obtenida en una cohorte y contexto académico específicos, no como prueba de generalización automática a otras asignaturas, programas o instituciones**.** Aunque los resultados muestran una capacidad predictiva alta en las ventanas analizadas, la validez externa del marco propuesto aún requiere contrastarse en nuevas cohortes del mismo curso, en otros cursos básicos de ingeniería y, posteriormente, en instituciones con estructuras curriculares, políticas de evaluación y niveles de actividad en el LMS diferentes. Estudios recientes han mostrado que el desempeño de los modelos puede variar notablemente entre contextos y cohortes, incluso cuando se emplean algoritmos similares [1], [5], [15], [17].

En segundo lugar, aunque se incorporaron variables académicas previas y de comportamiento, no se tuvieron en cuenta otras variables psicosociales (como la motivación, autoeficacia o la carga del trabajo externo a la universidad) que también se ha demostrado que influyen en la permanencia de los estudiantes. Para estudios posteriores puede ser interesante incorporar índices autorreportados junto con medidas objetivas de actividad para comprender mejor los riesgos y sus causas. En este sentido, la generalización del estudio debe entenderse en clave analítica y no universal. El aporte principal del trabajo no reside en afirmar que un mismo conjunto de predictores o umbrales funcionará de forma idéntica en cualquier entorno, sino en proponer una lógica metodológica replicable para construir sistemas de alerta temprana a nivel de asignatura. Variables como la estructura del curso, la frecuencia de evaluación, el tipo de actividades en el LMS, la proporción de estudiantes repetidores y el peso relativo de las evaluaciones parciales pueden modificar tanto la importancia de los predictores como el rendimiento del modelo. Por ello, la replicabilidad del marco exige ajustes contextuales y validaciones locales antes de su adopción operativa en otros escenarios.

En tercer lugar, el estudio no profundizó en detalle si el modelo puede estar sesgado respecto a ciertos subgrupos como el género, la trayectoria académica previa, o la modalidad de ingreso. En este sentido, los modelos explicativos interpretables de la equidad educativa muestran la importancia de evaluar los sistemas de alerta desde la perspectiva de la justicia y no discriminación.

Futuras líneas de investigación podrían centrarse en tres aspectos: (i) medir el verdadero impacto de las intervenciones ejecutadas a partir del sistema de alerta en las tasas de aprobado y retención, (ii) diseñar dashboards de analíticas que permitan la interactividad de docentes y alumnos con las evidencias, para explorar factores de riesgo y trayectorias de aprendizaje, y (iii) comparar sistemáticamente distintos esquemas de categorización del riesgo, considerando la carga del trabajo docente y el rendimiento académico obtenido. Adicionalmente, resulta prioritario desarrollar procesos de validación temporal y validación externa del modelo, replicando el análisis en nuevas cohortes del mismo curso y en otras asignaturas básicas de ingeniería impartidas en modalidad virtual. Este paso permitiría determinar hasta qué punto el marco mantiene su estabilidad predictiva cuando cambian las condiciones de evaluación, el perfil del alumnado y los patrones de interacción con el LMS.

VI. Conclusiones

Este estudio tuvo como propósito proponer y evaluar un marco de predicción temprana del riesgo de desaprobación en un curso básico de ingeniería impartido en modalidad virtual, a partir de analíticas de aprendizaje. Los resultados mostraron que los patrones tempranos de actividad en el LMS —en particular, los días activos, el número de accesos, la proporción de actividades evaluativas entregadas y la nota parcial acumulada— se asociaron de manera fuerte y consistente con el desenlace final de la asignatura, confirmando que el comportamiento de estudio observable en las primeras semanas constituye una señal robusta para la detección temprana del riesgo.

En primer lugar, se evidenció que las variables que aportaban más a la explicación del riesgo de desaprobación eran el porcentaje de entrega de tareas y cuestionarios, la nota acumulada en la semana 5, la regularidad de acceso al LMS y el GPA previo. Estas variables proporcionan a los docentes señales claras y tempranas sobre el compromiso y la trayectoria del estudiante, y convierten la conceptualización abstracta de "riesgo" en conductas concretas y observables.

En segundo lugar, los modelos supervisados alcanzan un alto rendimiento: la regresión logística obtiene AUC superiores a 0,80 en la ventana de la semana 5, mientras que los modelos en ensamble, como el random forest y el gradient boosting, se aproximan a un AUC de 0,90. Este nivel de discriminación coloca el marco propuesto dentro de un rango análogo al de la evidencia reciente en el ámbito de las analíticas de aprendizaje, sin renunciar a la interpretabilidad del modelo principal.

Por último, la conversión de las probabilidades de desaprobar en un score de riesgo y su categorización en niveles bajo, medio y alto permiten concentrar aproximadamente un 60 % de los estudiantes que finalmente desaprobaron en el grupo de nivel de riesgo más alto, interviniendo sólo en una cuarta parte de la cohorte. La relación entre la cobertura de los casos críticos y el esfuerzo de intervención da apoyo a la viabilidad del sistema de alerta temprana y proporciona a los equipos docentes un criterio concreto para priorizar recursos de apoyo académico en contextos de alta matrícula y limitaciones de tiempo.

Referencias

[1] P. West, F. Paige, W. Lee, N. Watts, y G. Scales, «Using Learning Analytics and Student Perceptions to Explore Student Interactions in an Online Construction Management Course», J. Civ. Eng. Educ., vol. 148, n.o 4, p. 05022001, oct. 2022, doi: 10.1061/(ASCE)EI.2643-9115.0000066.

[2] M. Rebelo Marcolino et al., «Student dropout prediction through machine learning optimization: insights from moodle log data», Sci Rep, vol. 15, n.o 1, p. 9840, mar. 2025, doi: 10.1038/s41598-025-93918-1.

[3] E.-Y. Seo, J. Yang, J.-E. Lee, y G. So, «Predictive modelling of student dropout risk: Practical insights from a South Korean distance university», Heliyon, vol. 10, n.o 11, p. e30960, jun. 2024, doi: 10.1016/j.heliyon.2024.e30960.

[4] J. B. Osborne y A. S. I. D. Lang, «Predictive Identification of At-Risk Students: Using Learning Management System Data», JPSS, vol. 2, n.o 4, pp. 108-126, jul. 2023, doi: 10.33009/fsop_jpss132082.

[5] M. Vaarma y H. Li, «Predicting student dropouts with machine learning: An empirical study in Finnish higher education», Technology in Society, vol. 76, p. 102474, mar. 2024, doi: 10.1016/j.techsoc.2024.102474.

[6] T. Ohkawauchi y E. Tanaka, «Predicting Student Dropout Risk Using LMS Logs», IIAI Letters on Institutional Research, vol. 4, p. 1, 2024, doi: 10.52731/lir.v004.226.

[7] O. O. Ajayi, «Predicting Student Dropout Risk in Online Learning Using Stacked Ensemble Machine Learning and Explainable AI Techniques», IJCA, vol. 187, n.o 40, pp. 26-29, sep. 2025, doi: 10.5120/ijca2025925707.

[8] M. Á. Rodríguez-Ortiz, P. C. Santana-Mancilla, y L. E. Anido-Rifón, «Machine Learning and Generative AI in Learning Analytics for Higher Education: A Systematic Review of Models, Trends, and Challenges», Applied Sciences, vol. 15, n.o 15, p. 8679, ago. 2025, doi: 10.3390/app15158679.

[9] P. Ngulube y M. M. Ncube, «Leveraging Learning Analytics to Improve the User Experience of Learning Management Systems in Higher Education Institutions», Information, vol. 16, n.o 5, p. 419, may 2025, doi: 10.3390/info16050419.

[10] M. Hernández-Campos, A. Gonzalez-Torres, y F. J. García-Peñalvo, «Learning Outcomes Evaluation Through Learning Analytics Systems in Higher Education: A Systematic Literature Review», SAGE Open, vol. 15, n.o 3, p. 21582440251347374, ene. 2025, doi: 10.1177/21582440251347374.

[11] F. Gutiérrez, K. Seipp, X. Ochoa, K. Chiluiza, T. De Laet, y K. Verbert, «LADA: A learning analytics dashboard for academic advising», Computers in Human Behavior, vol. 107, p. 105826, jun. 2020, doi: 10.1016/j.chb.2018.12.004.

[12] C. Lee et al., «Interpretable Predictive Analytics for Online Learning: A Markov-Based Machine Learning Approach», Learning Analytics, vol. 12, n.o 2, pp. 259-278, ago. 2025, doi: 10.18608/jla.2025.8375.

[13] K. Linden, N. Van Der Ploeg, y N. Roman, «Explainable learning analytics to identify disengaged students early in semester: an intervention supporting widening participation», Journal of Higher Education Policy and Management, vol. 45, n.o 6, pp. 626-640, nov. 2023, doi: 10.1080/1360080X.2023.2212418.

[14] J.-E. Russell, A. Smith, y R. Larsen, «Elements of Success: Supporting at-risk student resilience through learning analytics», Computers & Education, vol. 152, p. 103890, jul. 2020, doi: 10.1016/j.compedu.2020.103890.

[15] O. Goren, L. Cohen, y A. Rubinstein, «Early Prediction of Student Dropout in Higher Education using Machine Learning Models», en Proceedings of the 17th International Conference on Educational Data Mining, Atlanta, GA, USA: International Educational Data Mining Society, 2024, pp. 349-359. doi: 10.5281/ZENODO.12729831.

[16] J. M. F. Oro, P. G. Regodeseves, L. S. Bertolín, J. G. Pérez, R. Barrio-Perotti, y A. P. Blanco, «Application of Learning Analytics for the Study of the Virtual Campus Activity in an Undergraduate Fluid Mechanics' Course in Bachelors of Engineering», Tech Know Learn, vol. 30, n.o 3, pp. 1321-1344, sep. 2025, doi: 10.1007/s10758-024-09803-9.

[17] L. Paura, I. Arhipova, G. Vitols, y S. Sproge, «Analysis of Student Dropout Risk in Higher Education Using Proportional Hazards Model and Based on Entry Characteristics», Data, vol. 10, n.o 7, p. 110, jul. 2025, doi: 10.3390/data10070110.

[18] A. J. Pacheco, O. R. Boude Figueredo, A. Chiappe, y L. Fontán De Bedout, «AI-powered learning analytics for metacognitive and socioemotional development: a systematic review», Front. Educ., vol. 10, p. 1672901, sep. 2025, doi: 10.3389/feduc.2025.1672901.

[19] V. Čotić Poturić, S. Čandrlić, y I. Dražić, «A Scoring Algorithm for the Early Prediction of Academic Risk in STEM Courses», Algorithms, vol. 18, n.o 4, p. 177, mar. 2025, doi: 10.3390/a18040177.