

Impacto de la inteligencia artificial en el aprendizaje, modelo de predicción

Impact of artificial intelligence on learning: a prediction model

Silvia Soledad Moreno Gutiérrez¹, Luis Felipe Barrera Cervantes¹, Salvador Martín Camacho y López¹

¹ Universidad Autónoma del Estado de Hidalgo. Escuela Superior de Tlahuelilpan. Tlahuelilpan. Hgo. México. silviam@uaeh.edu.mx

Fecha de recepción: 22 de septiembre de 2025

Fecha de aceptación: 5 de marzo de 2026

Resumen. Durante la formación de estudiantes de nivel superior, en los últimos años el uso de las tecnologías inteligentes han representado una alternativa fundamental para el fortalecimiento del proceso educativo, de hecho, la evolución acelerada y el alto potencial de la inteligencia artificial (IA) están demandando la integración de docentes y estudiantes al ecosistema digital como una tendencia de aceptación creciente. Considerando este motor de innovación con miras hacia la sociedad 5.0, el rol de estudiante se está transformando, no así el rol del profesor quien muestra incertidumbre hacia la IA en diversos países de Latinoamérica. Por ello, el presente trabajo expone el desarrollo de 2 modelos de aprendizaje automático basados en la metodología CRISP-DM y datos de estudiantes de ingeniería de software, para analizar y predecir el impacto de IA en el aprendizaje. Luego de validación el mejor fue random forest con una precisión de 0.86, sensibilidad de 0.87 y f1 score de 0.84. La propuesta se considera adecuada para realizar la predicción y a través de sus resultados fortalecer la toma de decisiones y el proceso educativo mediante las herramientas pertinentes.

Palabras Clave: aprendizaje, aprendizaje automático, recursos inteligentes.

Summary. In recent years, the use of smart technologies has become a key alternative for strengthening the educational process in higher education; indeed, the rapid evolution and immense potential of artificial intelligence (AI) are driving the integration of both teachers and students into the digital ecosystem—a trend gaining increasing acceptance. While the student's role is evolving in light of this innovation engine and the move toward Society 5.0, the teacher's role remains static, often characterized by uncertainty regarding AI across various Latin American countries. Consequently, this study details the development of two machine learning models—based on the CRISP-DM methodology and data from software engineering students—to analyze and predict AI's impact on learning. Following validation, the Random Forest model emerged as the best performer, achieving an accuracy of 0.86, a sensitivity of 0.87, and an F1 score of 0.84. The proposed approach is deemed suitable for making predictions and leveraging the resulting insights to enhance decision-making and the educational process through appropriate tools.

Keywords: learning, machine learning, smart resources.

1 Introducción

En los últimos años la IA ha constituido una de las mayores tendencias mundiales y ha logrado transformar el ecosistema educativo, tanto en términos presenciales como virtuales el rol docente requiere sensibilización y atender las exigencias del estudiante que demanda nuevas competencias de sus facilitadores en términos disciplinares y didácticos [1].

Hoy en día, la innovación educativa inteligente es una necesidad para las escuelas de todos los niveles, más aun en el nivel profesional y posgrado donde los estudiantes se encuentran cerca o dentro del ámbito laboral. La presencia decisiva de la IA al ámbito educativo está obligando a la comunidad universitaria a ceñirse a una forma de trabajo acelerada e innovadora, en el caso de estudiantes esto ha sido bien recibido, no así por los docentes de diversos países de Latinoamérica quienes expresan desconocimiento y falta de credibilidad respecto a la IA como factor de éxito en la formación [2].

Países que, a su vez, reconocen la aplicación de alternativas inteligentes como factor decisivo para la adquisición del conocimiento y la superación de los desafíos actuales [3]. Según la revisión de la literatura, la implementación de la IA en el sector educativo ha generado debates acerca de su impacto en la enseñanza y el aprendizaje, diversos estudios han discutido las ventajas de los recursos educativos y las desventajas en el proceso según opinión de estudiantes y docentes [2]. El uso de estas aplicaciones ha mostrado en ocasiones

desconfianza, al igual que la capacitación docente en el tema [3], sin embargo, es fundamental la integración de recursos inteligentes y aprovechar el potencial de la IA [4].

Por lo anterior, la presente propuesta consiste en la creación de dos modelos predictivos del impacto de la IA en el aprendizaje cuyo propósito consiste en fundamentar la toma de decisiones alrededor del proceso educativo del estudiante desde un enfoque personalizado y con base en ello construir estrategias que contribuyan a incrementar su rendimiento a través de herramientas eficaces. Con base en los resultados se identificó la mejor propuesta, siendo random forest por encima de red neuronal artificial.

2 Estado del Arte

La literatura expresa la importancia de la IA en diversas actividades educativas, como es la planeación, impartición de clase y evaluación, sin embargo, se observan inconvenientes [5]. La IA podría contribuir a la personalización del aprendizaje mediante metodologías adecuadas, mediante ajuste dinámico y considerando la desigualdad en acceso tecnológico y la brecha educativa entre regiones [6], de igual manera podría apoyar la detección de deficiencias y construir estrategias efectivas, considerando limitaciones del alumno [7].

La literatura expone amplio número de trabajos en el ámbito de la educación apoyada por IA, sin embargo, son reducidas las propuestas para predecir impacto en el aprendizaje, por ello, se agregaron algunas por su representatividad en el ámbito.

Se desarrolló un predictor de rendimiento académico de alumnos de ingeniería donde se aplicaron técnicas de aprendizaje automático (AA), como son: K vecinos cercanos, máquina de soporte vectorial, árbol de decisión y perceptron simple logrando el mejor rendimiento con este último, no obstante, con precisión de 61% [8], en la predicción de resultado académico, aplicaron random forest y gradient boosted, logrando 90% y 93% de precisión [9]. Otros estudios analizaron regresión logística, xtreme boosting gradient, k-vecinos cercanos, random forest, red neuronal artificial y máquina de soporte vectorial, logrando la red neuronal un 82.4% de exactitud, siendo este el mejor resultado [10].

En la predicción del abandono escolar se desarrollaron 14 propuestas, siendo la de mayor precisión el K-vecino cercano con 87% de precisión [11].

3 Metodología

Para este apartado se expone un modelo de aprendizaje profundo para predecir el impacto de la IA en el aprendizaje de alumnos, para recabar los datos se aplicó un cuestionario mediante herramientas de google forms, logrando participación del 50% de la matrícula de ingeniería de software, es decir, 68 estudiantes. El trabajo se basó en la metodología CRISP-DM (Cross Industry Standard Process for Data Mining), la cual se integra por 6 fases: I. Entendimiento del negocio, del contexto, II. Entendimiento de los datos, III. Transformación de los datos, IV. Modelado, V. Validación y VI. Despliegue.

Fase I. La aplicación de recursos didácticos dentro del salón de clases en ocasiones se realiza de forma tradicional, es decir, los docentes en países latinos aun hacen uso de recursos diferentes a los que ofrece la IA, lo cual representa una debilidad reflejada en los estudiantes ubicados en desventaja respecto a aquellos que si aprovechan las tecnologías emergentes e innovadoras, quienes expresan la creciente demanda de experimentar experiencias educativas personalizadas y efectivas.

Fase II. Para identificar la percepción de los estudiantes respecto al impacto de la IA en el proceso de aprendizaje se realizó un análisis exhaustivo del comportamiento de los datos con la exploración de 21 variables, se contó con una variable objetivo la cual contiene 3 clases o categorías de salida. Las variables son de tipo numérico y object, (ver tabla 1). Para identificar cómo serán procesadas según el tipo de dato.

Tabla 1. Variables. Fuente: elaboración propia

VARIABLE	TIPO
Nivel educativo	Object
Has escuchado hablar sobre IA	Object

Herramienta de IA para fortalecer el aprendizaje	Object
Frecuencia de uso	Object
Tus profesores te permiten usar IA	Object
Actividades utilizan IA	Object
Actividades que la IA es útil para el aprendizaje	Object
La IA tiene desventajas para tu aprendizaje	Object
Principal desventaja	Object
La IA representa un peligro en la educación	Object
Beneficios del uso de la IA en el aprendizaje	Object
Uso responsable de la IA para el aprendizaje	Object
Es confiable la información generada por la IA	Int64
Impacto Aprendizaje	Int64
La IA mejorara la creatividad en los estudiantes	Object
La IA fomenta el pensamiento crítico	Object
Ha cambiado tus estudios con la IA	Object
A quién beneficia la IA en la educación	Object
Conoce la Sociedad 5.0	Object
Integrar IA en la enseñanza	Object
Evolución de IA en la educación	Object
Estrategias para la IA sin afectar el aprendizaje	Object

Fase III. La variable objetivo es ‘Impacto Aprendizaje’, aplica la siguiente función:

- Si el valor de Impacto Aprendizaje está entre 0 y 1 se clasifica como bajo.
- Si el valor de Impacto Aprendizaje está en 2, se clasifica como regular.
- Si el valor de Impacto Aprendizaje está entre 3 y 4, se clasifica como considerable.

Se procedió a verificar la existencia de valores nulos, lo que arrojó la existencia de ellos, por lo que se aplicó la moda para rellenar por tratarse de tipo categórico.

Con base en las variables procesadas por el modelo para su entrenamiento, se utilizó el método *Label Encoder* para la categorización. Para efectuar el escalamiento se utilizó la

técnica de *Standard Scaler*, que consiste alcanzar media=0 y desviación estándar=1. A continuación se procedió a verificar el equilibrio de los datos (ver figura 1).

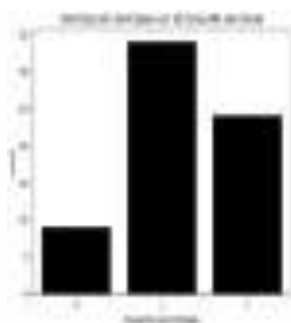


Figura 1. Equilibrio de los datos. Fuente: elaboración propia.

Los datos se dividieron en conjuntos de entrenamiento y prueba para evaluar el desempeño de los modelos. El proceso se realizó mediante la función *train test split* de la librería *Scikit-learn* con los siguientes parámetros: tamaño del conjunto de prueba: Test_size = 80%, semilla aleatoria: se estableció random_state = 42 para garantizar reproducibilidad.

Fase IV: El desarrollo de la propuesta consistió en un modelo de clasificación con algoritmo de Perceptrón Multicapa como un tipo de red neuronal que, en su arquitectura, ver figura 2, utiliza capas de nodos conectadas para aprender patrones ocultos.

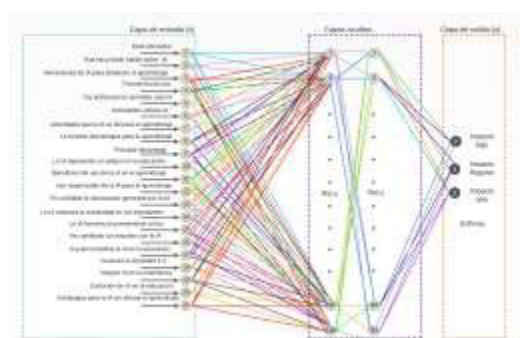


Figura 2. Arquitectura del modelo neuronal. Fuente: elaboración propia

De igual manera se aplicó el algoritmo Random Forest, el cual, se compone de una colección de árboles de decisión y cada árbol en el conjunto está compuesto por una muestra de datos existencia extraída por un conjunto de entrenamiento con reemplazo, cuya arquitectura consta de 800 árboles para procesar un total de 24 variables de entrada, correspondientes a las características del conjunto de datos, la figura 3 muestra su arquitectura.

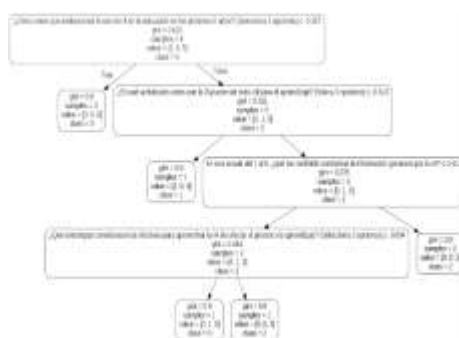


Figura 3. Arquitectura del modelo Random Forest. Fuente: elaboración propia

Después del entrenamiento se sometió a un conjunto de métricas y sus resultados son expuestos en el apartado siguiente.

4 Resultados

Fase V. Para la evaluación del modelo neuronal se analizó luego de equilibrar los datos con el fin de observar el comportamiento cada caso, el cual se muestra en la tabla 2.

Tabla 2. Resultados del modelo neuronal. Fuente: elaboración propia.

	Distribución original	Submuestreo	Sobremuestreo
Accuracy	0.62	0.51	0.59
Precisión	0.75	0.55	0.585
Recall	.665	0.55	0.60
f1-score	.665	0.55	0.58

Se graficó la pérdida y la precisión durante el entrenamiento (ver figura 4) para comprobar su comportamiento.

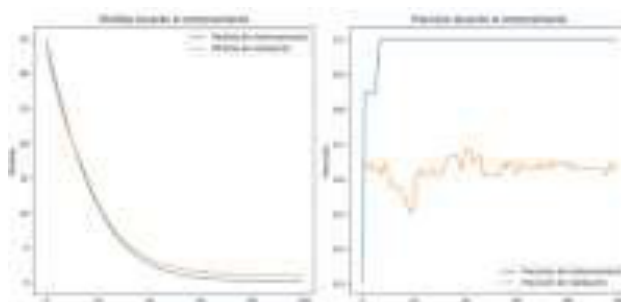


Figura 4. Pérdida y precisión modelo neuronal. Fuente: elaboración propia

Con base en lo anterior, la validación del algoritmo random forest se llevó a cabo utilizando las mismas métricas ya que al tratarse de un modelo de clasificación estas son adecuadas, sus resultados se exponen en la tabla 3.

Tabla 3. Resultados del modelo random forest. Fuente: elaboración propia.

Precisión	Recall	F1-score
0.86	0.87	0.84

Los resultados muestran que el rendimiento del modelo de Impacto_Aprendizaje es adecuado, lo cual representa alta confiabilidad en el uso. Las métricas aplicadas permitieron observar su desempeño luego de un proceso de validación, los resultados se visualizan de forma numérica. La percepción del impacto que tiene en el aprendizaje la IA es un fenómeno que algunos modelos lograron representar de manera favorable.

Fase VI. La aplicación y uso de los modelos se lleva a cabo como parte de proyecto de investigación, con el propósito de continuar el análisis del impacto que conlleva en los estudiantes, el uso de la IA y predecir escenarios posibles.

5 Conclusiones

El impacto y potencial de la IA en el aprendizaje es una pregunta fundamental que los docentes realizan de manera frecuente, esta incertidumbre podría dañar el aprendizaje y uso de tecnologías emergentes.

Se desarrolló un modelo neuronal y uno de random forest para el análisis del impacto de la IA en el aprendizaje y se obtuvieron resultados favorables del algoritmo de random forest, el cual se consideró adecuado.

Los modelos de AA muestran alto potencial para la representación de fenómenos cotidianos, tales como el aprendizaje del estudiante, no obstante, para predecir el problema que se expone el modelo neuronal demostró resultados deficientes.

En cuanto al análisis del impacto de la IA en el contexto académico, se considera adecuado construir modelos de AA híbridos, clasificación y regresión complementarios y orientados a la predicción y seguimiento de factores de éxito en las asignaturas, considerando la naturaleza diferente de cada una. El modelo propuesto es capaz de analizar patrones de comportamiento del estudiante y a través de ellos predecir el impacto de la IA en su aprendizaje.

La propuesta de predicción explota datos, no obstante, su privacidad es segura considerando que las herramientas de AA requieren la transformación de los datos antes de iniciar la construcción del modelo.

Como un trabajo futuro, se propone la construcción de modelos orientados a identificar deficiencias del estudiante y con base en ellas proponer métodos de aprendizaje y evaluaciones.

Referencias

- [1] Cárdenas, R. G. (2024). Educación superior e Inteligencia artificial: transitando de la docencia a la tutoría. <https://riesed.org/index.php/RIESED/article/view/175>

- [2] Tramallino, C. P., & Zeni, A. M. (2024). Avances y discusiones sobre el uso de inteligencia artificial (IA) en educación. *Educación*, 33(64), 29-54. <https://doi.org/10.18800/educacion.202401.m002>
- [3] Sánchez, O. V. G. (2023). Uso y percepción de ChatGPT en la educación superior. *Revista de Investigación En Tecnologías de la Información*, 11(23), 98-107. <https://doi.org/10.36825/riti.11.23.009>
- [4] Panaqué, C. R., & Castañón, C. B. (2024). La inteligencia artificial en la educación del siglo XXI: avances, desafíos y oportunidades. Presentación.
- [5] *Educación*, 33(64), 5-7. <https://doi.org/10.18800/educacion.202401.p001>
- [6] Montiel-Ruiz, F. J., & Ruiz, M. L. (2023). Inteligencia artificial como recurso docente en un colegio rural agrupado. *Revista Interuniversitaria de Investigación En Tecnología Educativa*, 28-40. <https://doi.org/10.6018/riite.592031>
- [7] Tandayamo, R. C. G., Haro, R. E. M., Lozada, R. F. L., Cobos, D. I. J., & Gaibor, W. A. C. (2023). La Inteligencia Artificial Utilizada como un Recurso para el Aprendizaje. *Dialnet*. <https://dialnet.unirioja.es/servlet/articulo?codigo=9146852>
- [8] Castaneda, A. U. (2023). Un viaje hacia la inteligencia artificial en la educación.
- [9] *Realidad y Reflexión*, 56, 121-136. <https://doi.org/10.5377/ryr.v1i56.15776>
- [10] Contreras, L. E., Fuentes, H. J., & Rodríguez, J. I. (2020). Predicción del rendimiento académico como indicador de éxito/fracaso de los estudiantes de ingeniería, mediante aprendizaje automático. *Formación universitaria*, 13(5), 233- 246.
- [11] Aráuz, D. C., & Martínez, J. J. (2023). Predicción del rendimiento académico en la UNADECA por medio de sistemas de clasificación. *Unaciencia Revista De Estudios E Investigaciones*, 16(31), 17-35.
- [12] López, T. S., Pérez, D. P., Ruíz, V. G., & Carvajal, L. F. (2023). Implantación de técnicas de machine learning y creación de una red neuronal artificial para la predicción del rendimiento académico de estudiantes en ambientes universitarios que usan e-learning y streaming. *DYNA*, 98(3), 282-287.
- [13] Rojas, J. P. J., Cañari, W. Y., & Galindez, J. A. O. (2024). Modelo de aprendizaje automático para la predicción del abandono escolar mediante encuestas de hogares. *Revista Identidad*, 10(2), 24-33.