




# Application of artificial intelligence to predict academic performance in order to prevent dropout in high school students: Systematic Literature Review

Salazar Alberto, Renzo Armando<sup>1</sup>, Guzman Aquije, Elvis Henry<sup>1</sup>, and Lozada Flores, Rose Mary<sup>1</sup>

<sup>1</sup>Universidad Tecnológica del Perú, Perú, U18202764@utp.edu.pe, C20203@utp.edu.pe, C16321@utp.edu.pe

*Abstract– This paper presents a systematic literature review on the application of artificial intelligence to predict student dropout in higher education. It highlights that student dropout not only affects individuals, but also institutional resources and society in general. To address this problem, the application of artificial intelligence was proposed as a tool to anticipate and enhance the academic performance of students. Through the identification of risk factors related to dropout and the provision of personalized intervention strategies. In doing so, the aim is to improve student retention and academic success. In the systematic literature review presented here, the PICO and PRISMA methodologies were used. For this purpose, Scopus databases were exhaustively searched and a total of 110 records were identified. After applying the inclusion and exclusion criteria, 18 academic articles were selected for review. The results of the systematic review indicate that the application of artificial intelligence can be effective in predicting student dropout and improving student retention. Different machine learning and deep learning models were found to be used to identify at-risk students and provide personalized recommendations. In addition, the importance of collecting and analyzing historical data to improve the accuracy of forecasting models was emphasized.*

*Keywords– Artificial intelligence, Academic performance, school dropout, Machine learning, Deep learning.*

# Aplicación de la inteligencia artificial para predecir el rendimiento académico con el fin de evitar el abandono estudiantil en alumnos de grado superior: Revisión Sistemática

Salazar Alberto, Renzo Armando<sup>1</sup>, Guzman Aquije, Elvis Henry<sup>1</sup>, and Lozada Flores, Rose Mary<sup>1</sup>

<sup>1</sup>Universidad Tecnológica del Perú, Perú, U18202764@utp.edu.pe, C20203@utp.edu.pe, C16321@utp.edu.pe

**Resumen**– Este documento presenta una revisión sistemática de literatura sobre la aplicación de la inteligencia artificial para predecir el abandono estudiantil en la educación superior. Donde se destaca que, el abandono de estudiantes no solo afecta a los individuos, sino también, a los recursos institucionales y a la sociedad en general. Para abordar este problema se propuso la aplicación de la inteligencia artificial como herramienta para anticipar y potenciar el desempeño académico de los estudiantes. A través de la identificación de los factores de riesgo relacionados con el abandono y la provisión de estrategias de intervención personalizadas. Con ello, se busca mejorar la retención estudiantil y el éxito académico. En la revisión sistemática de literature que se presenta se utilizaron las metodologías PICO y PRISMA. Para ello, se realizaron búsquedas exhaustivas en las bases de datos Scopus y se identificaron un total de 110 registros. Después de aplicar los criterios de inclusión y exclusión, se seleccionaron 18 artículos académicos para su revisión. Los resultados de la revisión sistemática indican que la aplicación de la inteligencia artificial puede ser efectiva para predecir el abandono estudiantil y mejorar la retención estudiantil. Se encontraron diferentes modelos de machine learning y Deep learning que se utilizaron para identificar a los estudiantes en riesgo y ofrecer recomendaciones personalizadas. Además, se destacó la importancia de la recopilación y análisis de datos históricos para mejorar la precisión de los modelos de predicción.

**Palabras clave**-- Inteligencia artificial, Rendimiento académico, abandono escolar, Aprendizaje automático, Aprendizaje profundo.

## I. INTRODUCCIÓN

En la actualidad la educación superior ha asumido un papel esencial en el desarrollo personal y profesional de los individuos, así como en el avance de las sociedades. Sin embargo, subsiste una problemática latente en este ámbito: el abandono de estudiantes. Esta situación no solo implica la pérdida de oportunidades para los estudiantes, sino también un mal uso de recursos institucionales y una preocupación a nivel global para los sistemas educativos. El sistema de evaluación de los estudiantes refleja el desempeño de los institutos de educación superior, lo que afecta el desarrollo de las generaciones futuras. Es una de las principales razones por las

que la mayoría de los países intentan hacer evolucionar el sistema educativo de forma continua [1].

Para hacer frente a este desafío, surge un interés creciente en la aplicación de inteligencia artificial como una herramienta para anticipar y potenciar el desempeño académico de los estudiantes. Esto se logra mediante la identificación de los factores de riesgo relacionados con el abandono y la provisión de estrategias de intervención personalizada.

El abandono de estudiantes en la educación superior es una problemática multifacética y compleja. Su impacto afecta tanto a instituciones educativas públicas como privadas, y sus determinantes pueden variar ampliamente, desde desafíos académicos y económicos hasta factores personales y psicológicos. Las consecuencias del abandono son significativas, ya que, no solo influyen en la trayectoria de los estudiantes, sino también, tienen implicaciones económicas y sociales tanto para las instituciones como para la sociedad en su conjunto. Por lo tanto, es esencial desarrollar herramientas eficaces para detectar de forma temprana a los estudiantes en riesgo de abandonar sus estudios y ofrecerles el apoyo necesario para alcanzar sus metas académicas [5].

La aplicación de inteligencia artificial en la predicción del abandono estudiantil y la mejora del rendimiento académico ofrece numerosos beneficios. Por un lado, permite analizar amplios conjuntos de datos recopilados por instituciones educativas, lo que, facilita la identificación de patrones y tendencias que podrían pasarse por alto con enfoques tradicionales. Por otro, posibilita la personalización de las estrategias de intervención, adaptándolas a las necesidades individuales de cada estudiante. Este enfoque no solo contribuye a prevenir el abandono, sino que, mejora la calidad general de la educación superior al proporcionar un respaldo más eficiente [3].

En este contexto se realizó una revisión de literatura sobre la aplicación de inteligencia artificial en proyectos relacionados con la predicción temprana del abandono de estudiantes en la educación superior con el fin de identificarlos y ofrecer recomendaciones personalizadas de retención. El objetivo de la investigación es profundizar la literatura científica en la implementación de modelos con el uso de machine learning y deep learning viabilizan el uso de redes

**Digital Object Identifier:** (only for full papers, inserted by LACCEI).

**ISSN, ISBN:** (to be inserted by LACCEI).

**DO NOT REMOVE**

neuronales artificiales. A partir de la revisión de artículos se analizó la propuesta de algoritmos de predicción para el aprendizaje automático más utilizados para el contexto de predecir el rendimiento académico y asegurar el éxito académico.

## II. METODOLOGÍA

Para la investigación se usó el enfoque de revisión sistemática, combinando la metodología PICO para identificar los objetos de estudio y PRISMA para la selección de artículos que se utilizaron en el análisis.

### A. Formulación de PICOC

De acuerdo con las recomendaciones de Richardson et al. [2] para una correcta formulación de preguntas, se debe tener en cuenta dos aspectos. En primer lugar, la interrogante debe estar completamente relacionada con el tema de estudio. En segundo lugar, la pregunta debe estar formulada de manera que la respuesta se obtenga de manera precisa. Los autores sugieren que, para lograr estos objetivos, se debe dividir la pregunta en cuatro componentes esenciales: (P) para referirse al paciente o población de interés, (I) para describir la intervención que se desea considerar, (C) para la comparación que se plantea y (O) para los resultados que interesa investigar.

Para aplicar este enfoque en el tema de investigación, se debe identificar cuidadosamente estos componentes en las distintas fases del proceso.

1) *Identificación de componentes PICOC:* En la etapa inicial, se identificó los elementos de estudio necesarios para la investigación que incluyen la Población de interés, la Intervención que se está considerando, la Comparación relevante, los Resultados de interés y el Contexto en el que se desarrolló el estudio, ver Tabla I.

TABLA I  
IDENTIFICACIÓN DE LOS COMPONENTES PICOC

<b>P</b>	Población	Estudiantes Universitarios
<b>I</b>	Intervención	Inteligencia artificial aplicada para la predicción del abandono de estudios
<b>C</b>	Comparación	Cuestionarios, Métodos tradicionales
<b>O</b>	Resultados	Medir rendimiento académico
<b>C</b>	Contexto	Abandono estudiantil

2) *Formulación de preguntas:* En la segunda fase, se usó los resultados de la metodología PICOC para formular una pregunta de investigación, el cual dio el resultado ¿Cuál es el impacto de la aplicación de la inteligencia artificial para

predecir el abandono estudiantil de alumnos de grado superior en base al rendimiento académico?

Para el proceso de búsqueda de artículos que se utilizaron en la investigación, se desglosó la pregunta en interrogantes específicas que se centraron en cada uno de los componentes de PICOC, ver Tabla II.

TABLA II  
PREGUNTAS PICOC

<b>P</b>	Población	Estudiantes Universitarios	¿Cuáles son los estudiantes que se han considerado para esta investigación?
<b>I</b>	Intervención	Inteligencia artificial aplicada para la predicción del rendimiento académico	¿Cómo se implementa la inteligencia artificial para predecir el rendimiento académico?
<b>C</b>	Comparación	Método tradicional	¿Qué métodos tradicionales se usan para evaluar el rendimiento académico?
<b>O</b>	Resultados	Medir rendimiento académico	¿Cómo se puede medir el rendimiento académico estudiantes?
<b>C</b>	Contexto	Abandono estudiantil	¿Cómo se puede reducir la tasa de abandono estudiantil de alumnos de grado superior?

3) *Identificación de palabras clave:* En la tercera etapa, se procedió a identificar las palabras clave relevantes para cada uno de los componentes presentes en la pregunta PICOC, ver Tabla III.

TABLA III  
PALABRAS CLAVE

<b>P</b>	Problema/Población	Estudiantes Universitarios	“University”, “students”, “At-risk students”, “Dropout students”, “Vulnerable students”, “Underrepresented students”
<b>I</b>	Intervención	Inteligencia artificial aplicada para la predicción del abandono de estudios	“Artificial Intelligence”, “IA”, “Prediction”, “Student attrition”, “undergraduate”, “educational institution”
<b>C</b>	Comparación	Método tradicional	“Survey”, “Questionnaire”, “bias”, “análisis of data”
<b>O</b>	Resultados	Medir rendimiento académico	“Academic Performance”, “Attendance”, “Assessment Scores”, “Course Completion”
<b>C</b>	Contexto	Abandono estudiantil	“Dropout”, “Higher Education”, “elite, social & productive development”.

4) *Implementación de operadores booleanos*: En la cuarta fase, se definió la estructura apropiada para llevar a cabo la búsqueda, utilizando el conector lógico "OR" y empleando comillas ("" ) para las palabras compuestas que forman parte de las palabras clave, ver Tabla IV.

TABLA IV  
SINTAXIS OR Y PALABRAS COMPUESTAS

<b>P</b>	Problema/Población	Estudiantes Universitarios	"University" OR "students" OR "At-risk students" OR "Dropout students" OR "Vulnerable students" OR "Underrepresented students"
<b>I</b>	Intervención	Inteligencia artificial aplicada para la predicción del abandono de estudios	"Artificial Intelligence" OR "IA" OR "Prediction" OR "Student attrition" OR "undergraduate" OR "educational institution"
<b>C</b>	Comparación	Método tradicional	"Survey" OR "Questionnaire" OR "bias" OR "analysis of data"
<b>O</b>	Resultados	Tasa de abandono de estudiantes	"Academic Performance" OR "Attendance" OR "Assessment Scores" OR "Course Completion"
<b>C</b>	Contexto	Abandono estudiantil	"Dropout" OR "Higher Education" OR "quality, social & productive development"

5) *Sintaxis de la fórmula PICO*: En la quinta y última fase, se estableció una relación entre los componentes PICOC utilizando el operador lógico "AND", ver Tabla V.

TABLA V  
ECUACIÓN DE BÚSQUEDA

<p>("University" OR "Students" OR "At-risk students" OR "Dropout students" OR "Vulnerable students" OR "Underrepresented students" AND "Artificial Intelligence" OR "IA" OR "Prediction" OR "Student attrition" OR "undergraduate" OR "educational institution" AND "Survey" OR "Questionnaire" OR "bias" OR "analysis of data" AND "Academic Performance" OR "Attendance" OR "Assessment Scores" OR "Course Completion" AND "Dropout" OR "Higher Education" OR "quality, social &amp; productive development")</p>
---

En esta fase, se obtuvieron un total de 422 resultados en la base de datos Scopus. Estos registros fueron sometidos a un

proceso de filtrado con el fin de seleccionar las fuentes que fueron utilizadas en la revisión sistemática de literatura.

6) Definición de criterios de inclusión y de exclusión

a) Criterios de Inclusión (CI):

- CI1: Los documentos incluidos deben usar metodologías basadas en inteligencia artificial.
- CI2: Los documentos deben basarse en predecir el rendimiento académico.
- CI3: Los documentos elegidos que sean artículos originales.
- CI4: Los artículos deben pertenecer al área de ingeniería.

b) Criterios de Exclusión (CE):

- CE1: Tipo de publicación que no tiene Acceso abierto.
- CE2: Publicaciones en idiomas diferentes al inglés y español.
- CE3: Documentos con antigüedad mayor a 5 años.
- CE4: Documentos que no se relacionaron con el tema.

### B. Prisma

En el proceso de selección de artículos para la investigación, se aplicó la metodología PRISMA, la cual ayudó a mejorar la calidad y la confiabilidad de las fuentes de información. Este proceso se dividió en 2 partes: la primera permitió identificar los artículos considerados para la inclusión, mientras que, la segunda, se utilizó para identificar los artículos que debían ser excluidos por diversos motivos con el fin de obtener un resultado de investigación de mayor Calidad.

1) Proceso de identificación: El proceso de selección de artículos constó de cinco etapas: identificación, duplicación, elegibilidad, selección y sesgo. En la etapa de identificación, se llevó a cabo una búsqueda exhaustiva de artículos en las bases de datos de Scopus. Estas bases de datos fueron seleccionadas debido a su amplia cobertura y la alta calidad y prestigio de las publicaciones que albergan. Ambas bases de datos implementan rigurosos procesos de selección y evaluación, lo que garantiza la calidad y confiabilidad de los artículos que contienen, respaldando así la validez de la información encontrada en ellas.

Posteriormente, se procedió a la identificación de los artículos relevantes para la investigación, resultando en un total de 110 registros.

2) Proceso de elegibilidad: La etapa de elegibilidad consistió en dos partes, en primer lugar, quitamos los registros no especializados en ingeniería, nos quedamos con 110 registros, y, en segundo lugar, excluimos registros que no tiene relación con la inteligencia artificial, quedaron 80 registros.

3) Proceso de selección: Primero se eliminó 30 documentos. Segundo, 12 que no eran artículos. Tercero, 23 por idioma. Cuarto, 16 por no encontrarse en el rango de los

años 2018 a 2023. Y, por último, 10 el resumen no guarda relación con el tema resultando 19.

4) Proceso de sesgo: Finalmente, se excluyó un artículo que tenía la página del Publisher caída, terminando con 18 artículos para la revisión sistemática, tal como se muestra en el siguiente diagrama de la Fig. 1.

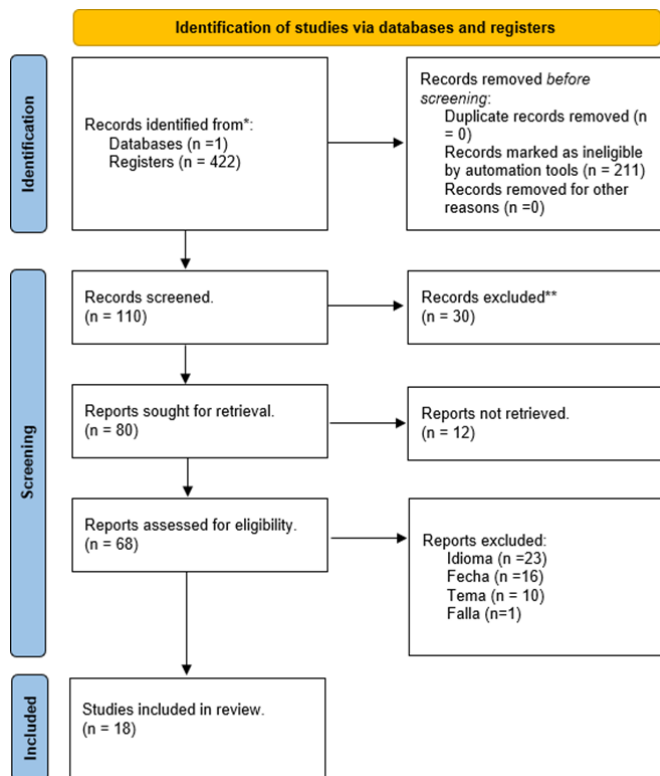


Fig. 1. Diagrama de flujo de PRISMA

### III. RESULTADOS

Con los resultados obtenidos se elaboró una tabla con el total de 18 artículos académicos, se asignó un código a cada uno de ellos para referenciarlos y recopilar información rápidamente. Luego, se elaboró un cuadro con preguntas específicas y a partir de la lectura de los artículos se respondió a las interrogantes formuladas.

En primer lugar, se identificaron las herramientas tecnológicas que los autores propusieron en sus investigaciones y se resumió en la Fig. 2.



Fig. 2. Metodologías eficientes

Se evidenció un claro predominio de Random Forest en las investigaciones, representado por 5 artículos revisados [3]–[8]. Le siguió la Red Neuronal Artificial (ANN) con 3 artículos revisados [1], [9], [10]. Por último, están el SOM-SIS con 1 artículo, la Regresión Lineal Múltiple, Random forest combinado con regresión lineal, Un enfoque híbrido, GA optimizado, Modelo Flexible, LSTM e IA vs Método tradicional [11]–[19]. Cabe señalar que, las herramientas Random Forest y la ANN fueron las más utilizadas porque generaron mejores resultados en las predicciones respecto al abandono estudiantil.

En la Fig. 3 se muestra el porcentaje de artículos de los años del 2020 al 2023 y se observó una mayoría de artículos publicados en los años 2022 y 2023, debido a la aparición de nuevas tendencias en la inteligencia artificial que surgen nuevas formas de pronóstico de la deserción estudiantil por medio del uso de algoritmos de predicción.

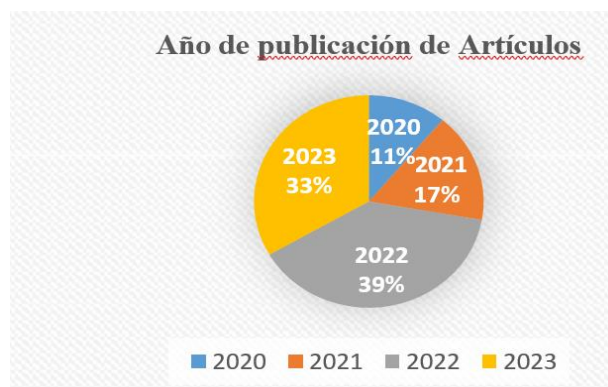


Fig. 3. Año de publicación de artículos

En la Fig. 4, se observa la cantidad de publicaciones que realizaron las editoriales donde destacó Universitas Ahmad Dahlan con 3 artículos. Esto, demostró un interés claro por encontrar las mejores herramientas para la disminución del abandono estudiantil.



Fig. 4. Editoriales de la publicación de Fuentes

En la fig. 5, se muestran los artículos por país. Del total de 19 artículos. La mayoría de los países cuenta solo con 1 artículo publicado. Sin embargo, en países como la India e Indonesia aportaron con 3 y 4 artículos respectivamente, se apreció un gran interés por la inteligencia artificial como herramienta para predecir el rendimiento académico y prevenir el abandono estudiantil.



Fig. 5. # de Artículos publicados por país

En la Fig. 6, se muestra la cantidad de fuentes por tipo de investigación. Donde se observó que existen más investigaciones cuantitativas, ya que, en la mayoría de los artículos se analizaron datos numéricos para comprender la relevancia de la inteligencia artificial en la generación de un pronóstico sobre el rendimiento académico para la toma de mejores decisiones.

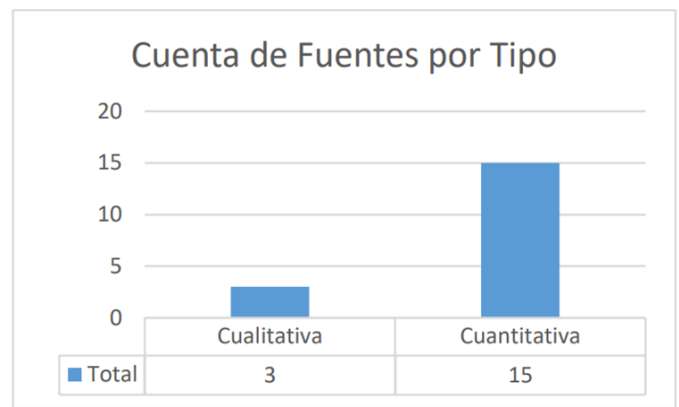


Fig. 6. Cantidad de fuentes por tipo de investigación

En la Tabla VI, se presentan las metodologías utilizadas por cada fuente y sus respectivos algoritmos que dejaron mejores resultados al momento de predecir el rendimiento académico de un estudiante.

TABLA VI  
METODOLOGÍAS

FUENTE	METODOLOGIA	MEJORES RESULTADOS
F1	Combinación de técnicas de mapa autoorganizado y técnica de preferenciade orden por similitudcon una solución ideal (SOM-SIS)	SOM - SIS
F2	Regresión lineal múltiple, regresión de vectores de soporte, árbol de decisión, bosque aleatorio, LassoCV y RidgeCV	regresión linealmúltiple
F3	La máquina de vectores de soporte, el bosque aleatorio, Naive Bayes, la regresión lógica y los árboles de decisión	Random Forest y de regresión logística
F4	Random Forest, Xtreme Boosting Gradient, Suport Vector Machine, K-NearestNeighbor y Logistic Regression, así como la creación de una red neuronal artificial (ANN)	Red neuronal artificial (ANN)
F5	Árbol de decisión (DT), regresión logística (LR), bosque aleatorio (RF), máquina de vectores de soporte (SVM), multicapa. Perceptrón (MLP), clasificadores Naive Bayes (NB), K-vecinos más cercanos (KNN), redes neuronales (NN)	un enfoque híbrido de aprendizaje automático que combina algoritmosde agrupación y clasificación

F6	Relief-F y Budget Tree-Random Forest	GA optimizado
F7	K vecino más cercano (KNN), máquina de vectores de soporte (SVM), análisis discriminante lineal (LDA) y bosque aleatorio (RF).	Random Forest
F8	J48, NBTree, OnerR, Decisión stump y Random Forest	Random Forest
F9	Método Tradicional	Modelo Flexible
F10	RF, SVM-RBF, K-NN, LR y NB	Random Forest
F11	Random forest, k vecinos más cercanos, árbol de decisión y red neuronal	Random Forest
F12	Deep Learning	LTSM
F13	Naive Bayes, Logistic Regression, Deep Learning, Random Forest, Decisión Tree y Gradient Boosted Tree	Random Forest
F14	ANCOVA, Logistic Regression, Support Vector Regression, Log-linear Regression, Decision Tree Regression, Random Forest Regression, and Partial Least Squares Regression	Log-linear Regression
F15	Log-linear Regression	Modelo de regresión múltiple
F16	Deep Learning	Red neuronal artificial (ANN)
F17	Deep Learning	Red neuronal artificial (ANN)
F18	IA	IA vs Método Tradicional

En la Fig. 7, se observa el tipo de herramienta tecnológica de la inteligencia artificial que fue utilizada en la recolección de fuentes académicas. Se obtuvo 12 artículos que refirieron a Machine Learning como la herramienta más utilizada, en la toma de mejores decisiones en base al análisis de datos realizado. 4 artículos de Deep Learning debido a su gran capacidad de adaptación y escalabilidad en el pronóstico de predicción del rendimiento académico. Otros 2 artículos, uno describió el uso de la inteligencia artificial vs método tradicional. El otro, explicó sobre el método tradicional basado en encuestas.

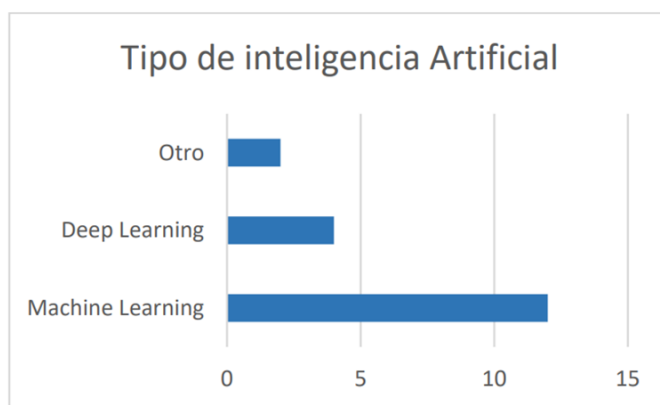


Fig. 7. Tipo de Inteligencia Artificial

#### IV. DISCUSIÓN

La educación superior se reconoce como crucial para el progreso social, proporcionando una variedad de ventajas que abarcan desde el logro de una carrera exitosa hasta la garantía de estabilidad financiera. Es por ello, que el foco de la investigación se centró en los estudiantes de educación superior.

La inteligencia artificial puede ser implementada a través de las herramientas de Machine Learning y el Deep Learning. Además, de tener variantes con las cuales se puede procesar datos de los alumnos y adecuar parámetros que se requieran [5].

Los métodos tradicionales usados para medir el rendimiento académico fueron estadísticas simples que tuvieron de los estudiantes, los cuales, no podían ser personalizadas ni redirigidas para aumentar su rendimiento [17].

Para medir el rendimiento académico se tuvo diferentes algoritmos como Decision Tree, Red Neuronal Artificial, Naive Bayer, Random Forest, SOM- SIS, Regresión lineal múltiple, entre muchas otras que ayudaron a predecir el rendimiento académico de un estudiante y tomar decisiones [9].

Entre las opciones más eficientes para medir la tasa de abandono tuvieron: Random Forest y las Redes neuronales artificiales que según [5] y [3] sustentaron que fueron la mejor opción para medir el rendimiento académico y predecir la deserción estudiantil en alumnos de grado superior.

#### V. CONCLUSIONES

El uso de la inteligencia artificial para anticipar el abandono de estudiantes en la educación superior es un asunto que ha cobrado importancia y suscita creciente interés. La capacidad de detectar de manera precoz a los estudiantes en situación de riesgo brinda a las instituciones educativas la oportunidad de implementar intervenciones personalizadas y efectivas, con el objetivo de mejorar la retención estudiantil y el desempeño académico.

Añadiría una propuesta de combinar un algoritmo de regresión lineal múltiple para la recomendación de grupos

para el mismo perfil de estudiante académico junto con un algoritmo de Random Forest para predicción de rendimiento académico, lo cual permitiría recuperar a los alumnos que estén mal académicamente y reinsertarlos en grupos de estudios con estudiantes de su mismo perfil con el fin de lograr una retroalimentación conjunta y conseguir aumentar su rendimiento académico.

Las redes neuronales artificiales (RNA) son recursos significativos en el análisis de datos, empleados para descubrir y expresar correlaciones funcionales entre variables. En la actualidad, las RNA se emplean para enfrentar, prever y categorizar desafíos en distintas disciplinas. Un ejemplo de su aplicación se encuentra en el entorno académico, donde las RNA se utilizan para evaluar el rendimiento estudiantil [10]. Por lo que, es un algoritmo con una base sólida para realizar una mejora del rendimiento académico en estudiantes.

La inteligencia artificial, a través de técnicas como Machine Learning y Deep Learning han demostrado ser herramientas efectivas para predecir el rendimiento académico y mejorar la retención estudiantil en la educación superior. La combinación de algoritmos como regresión lineal múltiple y Random Forest, con el uso de Redes Neuronales Artificiales vía el entrenamiento de árboles de decisión generaron el pronóstico que mejoró la toma de decisiones en la gestión de la deserción estudiantil, y con ello, promovió el número de estudiantes que continuaron sus estudios o terminaron su carrera. Estos hallazgos subrayaron la relevancia de la inteligencia artificial en la mejora continua de la calidad educativa en instituciones de educación superior.

#### REFERENCIAS

- [1] K. Borhani and R. T. K. Wong, "An artificial neural network for exploring the relationship between learning activities and students' performance". *Decision Analytics Journal*, vol. 9, 2023, doi: 10.1016/j.dajour.2023.100332.
- [2] S. Richardson W, M. C. Wilson, J. Nishikawa, and R. Hayward, "The well-built clinical question: a key to evidence-based decisions". *ACP J Club*, vol. 123, no. 3, pp. 12–13, Nov. 1995.
- [3] E. Ismanto, H. Ab. Ghani, and N. I. B. M. Saleh, "A comparative study of machine learning algorithms for virtual learning environment performance prediction". *IAES International Journal of Artificial Intelligence*, vol. 12, no. 4, pp. 1677–1686, 2023, doi: 10.11591/ijai.v12.i4.pp1677-1686.
- [4] S. Begum and S. S. Padmannavar, "Genetically Optimized Ensemble Classifiers for Multiclass Student Performance Prediction". *International Journal of Intelligent Engineering and Systems*, vol. 15, no. 2, pp. 316–328, 2022, doi:10.22266/ijies2022.0430.29.
- [5] P. J. A. L. Rose and T. A. C. Mary, "An Early Intervention Technique for At-Risk Prediction of Higher Education Students in Cloud-based Virtual Learning Environment using Classification Algorithms during COVID-19". *International Journal of Advanced Computer Science and Applications*, vol. 13, no. 1, pp. 612–621, 2022, doi: 10.14569/IJACSA.2022.0130174.
- [6] E. Mushtaha, S. Abu Dabous, I. Alsyouf, A. Ahmed, and N. Raafat Abdraboh, "The challenges and opportunities of online learning and teaching at engineering and theoretical colleges during the pandemic". *Ain Shams Engineering Journal*, vol. 13, no. 6, 2022, doi: 10.1016/j.asej.2022.101770.
- [7] M. N. Injadat, A. Moubayed, A. B. Nassif, and A. Shami, "Multi-split optimized bagging ensemble model selection for multi-class educational data mining". *Applied Intelligence*, vol. 50, no. 12, pp.4506–4528, 2020, doi: 10.1007/s10489-020-01776-3.
- [8] A. Pejić and P. S. Molcer, "Predictive machine learning approach for complex problem-solving process data mining". *Acta Polytechnica Hungarica*, vol. 18, no. 1, pp. 45–63, 2021, doi: 10.12700/APH.18.1.2021.1.4.
- [9] T. Santamaría-Lopez, D. Patiño-Perez, V. González- Ruiz, and L. Flores-Carvajal, "Implementation of machine learning techniques and creation of an artificial neural network for the prediction of the academic performance of students in university environments that use e-learning and streaming". *Dyna (Spain)*, vol. 98, no. 3, pp. 282–287, 2023, doi: 10.6036/10760.
- [10] Y. Baashar et al., "Toward Predicting Student's Academic Performance Using Artificial Neural Networks (ANNs)". *Applied Sciences (Switzerland)*, vol. 12, no. 3, 2022, doi: 10.3390/app12031289.
- [11] F. Ouyang, M. Wu, L. Zheng, L. Zhang, and P. Jiao, "Integration of artificial intelligence performance prediction and learning analytics to improve student learning in online engineering course". *International Journal of Educational Technology in Higher Education*, vol. 20, no. 1, 2023, doi: 10.1186/s41239-022-00372-4.
- [12] K. Fahd, S. J. Miah, and K. Ahmed, "Predicting student performance in a blended learning environment using learning management system interaction data". *Applied Computing and Informatics*, 2021, doi: 10.1108/ACI-06-2021-0150.
- [13] Y. Muhammad, M. A. Hassan, S. Almotairi, K. Farooq, F. Granelli, and E. Strážovská, "The Role of Socioeconomic Factors in Improving the Performance of Students Based on Intelligent Computational Approaches". *Electronics (Switzerland)*, vol. 12, no. 9, 2023, doi: 10.3390/electronics12091982.
- [14] K. R. Kannan, K. T. M. Abarna, and S. Vairachilai, "Open Issues, Research Challenges, and Survey on Education Sector in India and Exploring Machine Learning Algorithm to Mitigate These Challenges". *International Journal on Recent and Innovation Trends in Computing and Communication*, vol. 11, no. 6s, pp. 283–288, 2023, doi: 10.17762/ijritcc.v11i6s.6931.
- [15] R. O. Aluko, E. I. Daniel, O. Shamsideen Oshodi, C.O. Aigbavboa, and A. O. Abisuga, "Towards reliable prediction of academic performance of architecture students using data mining techniques". *Journal of Engineering, Design and Technology*, vol. 16, no. 3, pp. 385–397, 2018, doi: 10.1108/JEDT-08-2017-0081.
- [16] E. Ismanto, H. A. Ghani, N. I. M. Saleh, J. Al Amien, and R. Gunawan, "Recent systematic review on student performance prediction using backpropagation algorithms". *Telkomnika (Telecommunication Computing Electronics and Control)*, vol. 20, no. 3, pp. 597–606, 2022, doi: 10.12928/TELKOMNIKA.v20i3.21963.
- [17] I. E. Guabassi, Z. Bousalem, R. Marah, and A. Qazdar, "Comparative Analysis of Supervised Machine Learning Algorithms to Build a Predictive Model for Evaluating Students' Performance". *International journal of online and biomedical engineering*, vol. 17, no. 2, pp. 90–105, 2021, doi: 10.3991/ijoe.v17i02.20025.
- [18] A. A. Alnawas, M. M. H. Al-Jawad, and H. Alharbi, "A prediction model based on students's behavior in e-learning environments using data mining techniques". *Journal of Engineering and Sustainable Development*, vol. 26, no. 5, pp. 115–126, 2022, doi: 10.31272/jeasd.26.5.11.
- [19] R. Widayanti, S. Madenda, E. P. Wibowo, and K. Anwar, "SOM-SIS approach to auto summary of clustering results on university academic performance". *Telkomnika (Telecommunication Computing Electronics and Control)*, vol. 21, no. 1, pp. 104–112, 2023, doi: 10.12928/TELKOMNIKA.v21i1.24238.