

Inteligencia artificial para analizar el rendimiento académico en instituciones de educación superior. Una revisión sistemática de la literatura
Artificial Intelligence for analyzing academic performance in higher education institutions. A systematic literature review

Patricia Jimbo-Santana¹, 0000-0001-7432-1622

Laura Cristina Lanzarini², 0000- 0001- 7027-7564

Mónica Jimbo-Santana³, 0000-0002-3948-9507

Mario Morales-Morales⁴, 0000-0002-7493-8072

¹ Universidad Central del Ecuador, Quito, Ecuador. Facultad de Ciencias Administrativas.

prjimbo@uce.edu.ec

² Universidad Nacional de La Plata, La Plata, Argentina. Facultad de Informática.

laural@lidi.info.unlp.edu.ar

³ Universidad Central del Ecuador, Quito, Ecuador. Facultad de Ciencias Administrativas.

djimbo@uce.edu.ec

⁴ Universidad Central del Ecuador, Quito, Ecuador. Facultad de Ingeniería y Ciencias Aplicadas.

mmoralesm@uce.edu.ec.

Resumen

La inteligencia artificial se encuentra en constante evolución y se está aplicando en varias áreas, dentro de esta se encuentra la educación. El análisis del rendimiento académico de los estudiantes en instituciones de educación superior es un tema crítico para la toma de decisiones y el mejoramiento de la calidad de la educación. El objetivo de este artículo es realizar una revisión sistemática de la literatura, considerando las investigaciones que se han desarrollado utilizando técnicas de inteligencia artificial para analizar el rendimiento académico en instituciones de educación superior. Se han considerado las bases de datos científicas Web of Science, Scopus, e IEEE Xplore. Se consideró palabras clave relacionadas con la inteligencia artificial y el rendimiento académico. Se tomó en cuenta artículos publicados desde enero 2017 hasta diciembre 2022, se obtuvieron 1427 manuscritos, de los cuales se seleccionaron y analizaron 74, según los criterios de inclusión y exclusión predefinidos. Entre los resultados obtenidos se puede indicar que las técnicas más utilizadas para la predicción del rendimiento académico son: las redes neuronales y los árboles de decisión. En conclusión, se puede indicar que la aplicación de la inteligencia artificial puede mejorar la eficiencia y precisión de la evaluación, y proporcionar información valiosa para la toma de decisiones y la mejora de la calidad de la educación. Adicionalmente, se discuten las implicaciones y limitaciones de estos estudios y se proponen áreas para futuras investigaciones.

Palabras clave: Educación superior, inteligencia artificial, rendimiento académico, revisión sistemática.

Abstract

Artificial intelligence is constantly evolving and is being applied in several areas, including education. The analysis of the academic performance of students in higher education institutions is a critical issue for decision making and improving the quality of education. The objective of this article is to perform a systematic review of the literature, considering the research that has been developed using artificial intelligence techniques to analyze academic performance in higher education institutions. The scientific databases Web of Science, Scopus, and IEEE Xplore were considered. Keywords related to artificial intelligence and academic performance were considered. Articles published from January 2017 to December 2022 were taken into account, 1427 manuscripts were obtained, from which 74 were selected and analyzed, according to the predefined inclusion and exclusion criteria. Among the results obtained, it can be indicated that the most used techniques for the prediction of academic performance are: neural networks and decision trees. In conclusion, it can be indicated that the application of artificial intelligence can improve the efficiency and accuracy of the evaluation, and provide valuable information for decision making and improvement of the quality of

education. In addition, the implications and limitations of these studies are discussed and areas for future research are proposed.

Keywords: higher education, artificial intelligence, academic performance, systematic review.

Recibido: 17/04/2023

Aceptado: 26/05/2023

1. Introducción

En la actualidad estamos inmersos en una sociedad cada vez más sujeta a procesos tecnológicos a gran escala:

El área de la educación (que es sensible a los cambios en la sociedad ya que avanza a la par de esta) también se encuentra atravesando dicha tendencia de adaptación a las novedosas comunidades de interacción tecnológica; proceso que está orientado a nuevas tendencias y perfiles en relación con las nuevas propuestas en el sector (Ocaña, 2019, p. 537).

La inteligencia artificial representa un sinónimo de innovación y tecnología, ya que a partir de ella se imitan los procesos de inteligencia humana con el uso de algoritmos y ayuda a encontrar patrones de comportamiento que no son visibles a simple vista. En el ámbito de la educación ha representado gran ayuda en varios campos, dentro de estos la predicción del rendimiento académico, gracias a que la inteligencia artificial es la encargada del procesamiento de gran cantidad de datos, lo que nos permite entender de mejor manera los perfiles de los estudiantes y con ello sus necesidades. Con base en estos requerimientos se pueden establecer planes, proyectos y actividades académicas que ayuden a mejorar la calidad de la educación, comprender el perfil de los estudiantes y sus necesidades en cuanto al nivel de progreso o inicios de deserción, lo que resulta importante ya que de esta forma las instituciones pueden tomar decisiones que brinden apoyo en los procesos de aprendizaje, y a la vez mantengan un seguimiento constante del nivel de calidad de la educación.

Si se hace referencia a teorías educativas se puede mencionar la teoría del diseño instruccional que utiliza datos recopilados a través de la tecnología digital y la inteligencia artificial para informar el diseño de entornos de aprendizaje efectivos. Se ha observado que la inteligencia artificial es capaz de analizar datos sobre el rendimiento académico de los estudiantes y proporcionar información valiosa para mejorar el diseño de materiales, actividades y evaluaciones. Otra teoría relevante es la del aprendizaje adaptativo, la cual propone adaptar la instrucción y los recursos educativos según las necesidades individuales de cada estudiante. Mediante el uso de algoritmos de aprendizaje automático, la inteligencia artificial puede analizar datos sobre el rendimiento académico y ajustar la entrega de contenido, la secuencia de actividades y las estrategias de enseñanza con el objetivo de optimizar el aprendizaje de cada estudiante.

Las aplicaciones basadas en inteligencia artificial para la educación han crecido rápidamente en todo el mundo, y han afectado a las instituciones de educación superior. Varias de estas aplicaciones se están implementando para mejorar los servicios universitarios, ayudar a los docentes a ofrecer una educación de calidad, tomar decisiones en cuanto a la deserción universitaria y apoyar el aprendizaje de los estudiantes (Salas-Pilco y Yang, 2022).

Las instituciones de educación superior disponen de información detallada de cada uno de los estudiantes, sin embargo, carecen de modelos que les permitan describir al estudiante de manera objetiva, y con base en este análisis tomar decisiones que les permitan evitar problemas como por ejemplo la deserción estudiantil (Baldino et al., 2016, p. 589).

Por otra parte, a través del tiempo se ha podido evidenciar un constante crecimiento de las aplicaciones de inteligencia artificial en los procesos educativos, siendo uno de los campos emergentes la tecnología educativa, pues si bien ha existido durante aproximadamente 30 años, todavía no está claro si se puede impactar significativamente en la enseñanza y el aprendizaje en la educación superior (Zawacki-Richter et al., 2019, p. 2).

Algunos de los problemas que se pueden presentar en las instituciones de educación superior al no utilizar la inteligencia artificial en la evaluación del rendimiento académico pueden ser: falta de eficiencia en la evaluación, dificultad para identificar patrones y tendencias, falta de intervención temprana en el rendimiento de los estudiantes, limitaciones en la toma de decisiones institucionales, entre otras. No se debe dejar de lado que el sistema educativo está pasando de centrarse en el producto al proceso, expandiéndose más allá del

dominio del conocimiento para incluir la autorregulación, la colaboración y la motivación (Roll y Wylie, 2016).

En las investigaciones que utilizan inteligencia artificial para predecir el rendimiento académico, se utilizan diversas variables para construir modelos de predicción, como el historial académico del estudiante, su edad, género, nivel socioeconómico, entre otros. Estas variables son procesadas por algoritmos de aprendizaje automático que buscan identificar patrones y relaciones entre ellas, y de esta manera predecir el rendimiento académico futuro del estudiante.

En la actualidad, el no emplear técnicas de inteligencia artificial en el rendimiento académico podría implicar una falta de eficiencia en la gestión de la información, crear sesgos y subjetividades, así como la falta de personalización y adaptación a las necesidades específicas de los estudiantes.

El objetivo del presente trabajo es realizar una revisión sistemática de la literatura que considere las investigaciones realizadas desde enero 2017 hasta diciembre 2022, sobre las técnicas de inteligencia artificial (minería de datos, deep learning, big data), para analizar el rendimiento académico en instituciones universitarias, considerando como el resultado obtenido por un estudiante en una o varias pruebas estandarizadas, como exámenes, tareas, proyectos, entre otros, que miden el nivel de conocimiento y habilidades adquiridos en un determinado periodo de tiempo. El rendimiento académico también puede ser medido en función de la asistencia a clases, la participación en actividades extracurriculares, el cumplimiento de tareas y responsabilidades escolares, el desempeño en trabajos en equipo y otros indicadores que reflejen el desempeño académico del estudiante.

Este trabajo está organizado de la siguiente forma: la sección 2 describe los materiales y métodos utilizados, la sección 3 indica el análisis realizado y los resultados obtenidos, la sección 4 se encarga de la Discusión, y finalmente la sección 5 indica las conclusiones y trabajos futuros.

2. Materiales y métodos

Es importante primero considerar los objetivos de esta investigación:

- Analizar la evolución de las técnicas de inteligencia artificial aplicadas al rendimiento académico en la educación superior.
- Identificar los países en los que se han desarrollado estas investigaciones.
- Evaluar el tamaño de la muestra utilizado en las investigaciones sobre el análisis del rendimiento académico en las instituciones de educación superior.
- Identificar y describir las técnicas de inteligencia artificial utilizadas para el análisis del rendimiento académico en las instituciones de educación superior.

2.1. Método

La investigación utilizada es una revisión sistemática de la literatura, que tiene como objetivo sintetizar y analizar críticamente la evidencia disponible en un tema específico. Esto implica realizar un conjunto de actividades enmarcadas en tres fases, i) planteamiento, ii) conducción, y, iii) presentación, siguiendo el esquema propuesto por (Kitchenham, 2014) y que da sustento formal a esta investigación (Figura 1). Para ello, se realiza una búsqueda sistemática y exhaustiva en bases de datos y otras fuentes relevantes, se aplican criterios de inclusión y exclusión predefinidos y se analizan los estudios seleccionados.

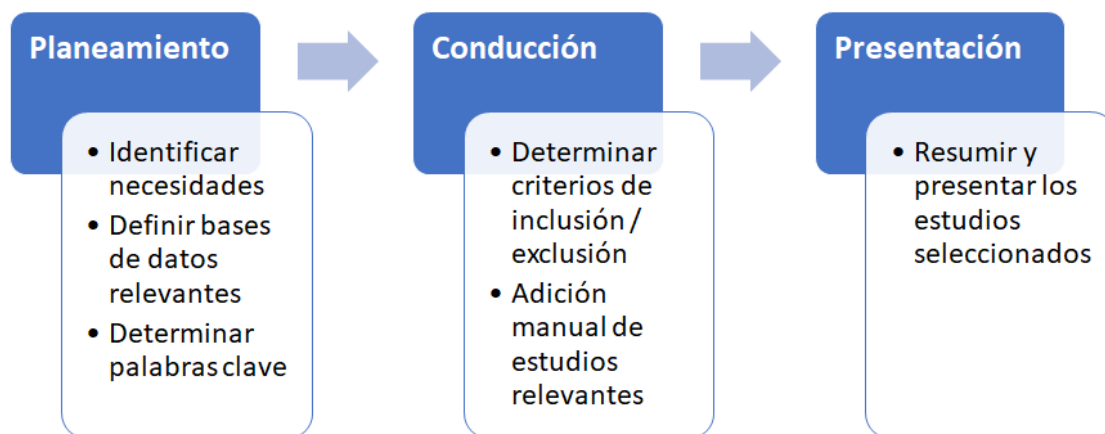


Figura 1: Actividades para la revisión sistemática

Las técnicas y estrategias para el estudio incluyeron la búsqueda en las bases de datos científicas Web of Science, Scopus y IEEE Xplore, con palabras clave relacionadas con inteligencia artificial y rendimiento académico. Se seleccionaron artículos publicados desde enero de 2017 hasta diciembre de 2022 y se aplicaron criterios de inclusión y exclusión predefinidos para seleccionar los 74 artículos analizados. Además, se utilizó un enfoque crítico y analítico para sintetizar los resultados y discutir las implicaciones y limitaciones de los estudios encontrados.

Según Brereton (2007) “realizar una revisión sistemática implica varias actividades discretas, que se pueden agrupar en tres fases principales: planificación, realización de la revisión e informe de la revisión” (p. 2), lo que se considera para determinar la situación actual de la inteligencia artificial, en relación con el rendimiento académico en la educación superior.

En la ejecución del protocolo se identificó la necesidad de la revisión, se determinaron las preguntas que van a guiar la investigación, se identificaron los términos, estrategias y las cadenas de búsqueda (inglés y español), se establecieron los criterios de inclusión para proceder a la extracción de datos. Finalmente se ejecutó y se desarrolló el protocolo de revisión.

2.2. Necesidad de la revisión

Realizando la respectiva revisión, se encontraron varios artículos de revisión sistemática sobre la utilización de la inteligencia artificial en la educación superior, enfocados en varios campos, así podemos indicar:

Chiu (2023) realiza una revisión sistemática entre 2012 – 2021 “haciendo énfasis en cuatro dominios: el aprendizaje, la enseñanza, la evaluación y la administración” (p. 1). En otro estudio de realizaron una revisión sistemática entre 2012 – 2020 donde se enfocaron en el aprendizaje en línea, el mismo que se ha implementado de poco a poco en la educación superior en los últimos años, además de traer consigo desafíos y nuevos retos, también permite a las instituciones mejorar la instrucción y el aprendizaje (Ouyang et al., 2022, p. 23).

También se realizaron una revisión de literatura, analizando artículos desde 1980 hasta el 2019, aquí mencionan el avance tecnológico que trae la revolución Industrial 4.0, esto permite que las instituciones de educación superior se vean obligadas a enfrentar la transformación digital y considerar dentro de ello a la inteligencia artificial (Castro-Benavides et al., 2020, p.1). En otro trabajo de realizan un marco de referencia teórico, analizando la transformación digital en las instituciones de educación superior (Almaraz Menéndez et al., 2017).

Otras investigaciones se enfocan en los cambios que se han generado en las instituciones de educación superior, esto debido a la inteligencia artificial, además de los nuevos retos que vienen de la mano con esta, concluyendo así que la inteligencia artificial es un desafío de gran importancia para las Instituciones de educación superior, pero que sin embargo en el caso de América Latina esta realidad resulta todavía inalcanzable (Ocaña Fernández, et al., 2019, p. 8). Otra investigación se centra en las oportunidades, riesgos y posibilidades a las que se enfrenta la educación, y las ventajas que se obtienen al utilizar la inteligencia artificial (Barrios Tao et al., 2021, p. 1). Otro estudio habla sobre la evolución que se ha evidenciado en la tecnología, y de cómo la inteligencia artificial se ha aplicado en diferentes ámbitos en las instituciones de educación superior (Ouyang y Jiao, 2021).

De estas investigaciones, ninguna de ellas aporta información significativa en el área del rendimiento académico en las instituciones de educación superior. Es importante considerar que la tecnología avanza día a día, por lo que es necesario realizar una revisión de la literatura que incorpore las últimas investigaciones.

2.3. Preguntas de Investigación

Las preguntas de investigación que guiaron la revisión sistemática son:

- ¿Cómo han evolucionado las técnicas de inteligencia artificial aplicadas al rendimiento académico en la educación superior?
- ¿En qué países se han desarrollado estas investigaciones?
- ¿Cuál es el tamaño de la muestra que se ha utilizado en las investigaciones?
- ¿Qué técnicas de inteligencia artificial han sido utilizadas para análisis del rendimiento académico en las instituciones de educación superior?

2.4. Términos de búsqueda

A partir de un sondeo preliminar de literatura se identificaron los términos de búsqueda que están contenidos en el título, las palabras clave y el resumen de los artículos a través de distintas combinaciones de términos principales y secundarios, los mismos que se indican en el Cuadro 1, se utilizó operadores lógicos AND y OR. La cadena de búsqueda en inglés que se utilizó fue:

(“intelligence artificial” OR AI) AND (“Academic Performance” OR “College Performance”) AND («higher education» OR university OR college).

Cuadro 1: Términos de Búsqueda

Términos principales en español	Términos alternativos en español	Términos principales en inglés	Términos alternativos en inglés
Inteligencia Artificial	IA	Artificial Intelligence	AI
Rendimiento académico	Desempeño universitario	Academic Performance	College Performance
Educación superior	Universidad	Higher Education	University College

2.5. Estrategias de búsqueda

Las bases de datos seleccionadas incluyeron Web of Science, Scopus, e IEEE Xplore, estas se seleccionaron por su relevancia en esta línea de investigación.

2.6. Criterios de inclusión

Los artículos fueron seleccionados de acuerdo con los siguientes criterios:

Cuadro 2: Criterios de inclusión

Criterios de inclusión	Criterios de exclusión
------------------------	------------------------

-
- Los artículos científicos deben tener relación de la aplicación de la inteligencia artificial para predecir y/o describir el rendimiento académico de los estudiantes en las instituciones de educación superior.
 - Los artículos científicos deben tener relación de la aplicación de la inteligencia artificial para predecir y/o describir el rendimiento académico de los estudiantes en las instituciones de educación superior.
 - Los estudios o artículos científicos son relevantes para responder a las preguntas de investigación.
 - Los artículos científicos deben abarcar todo el ámbito universitario, es decir, sin distinción de país., universidad. Carreras o profesiones.
 - Los estudios o artículos científicos son relevantes para responder a las preguntas de investigación.
 - Los estudios deben estar dentro del rango de investigación desde enero 2017 a diciembre 2022.
 - Los estudios deben estar en idioma inglés y español.
 - Artículos que incluyan las cadenas de búsqueda en el título, palabras claves, o en el resumen.
 - Artículos publicados en revistas o conferencias
 - Artículos científicos disponibles a texto completo.
 - Estudios que aplicaban la inteligencia artificial para predecir o pronosticar el rendimiento académico en la educación en escuelas, colegios y jardines.
 - Artículos científicos que se centran en herramientas de inteligencia artificial para el proceso de enseñanza aprendizaje, ingeniería, educación física, programación, entre otros.
 - Estudios realizados antes del 2017.
-

2.7. Extracción de datos

Se estructuraron varias cadenas de búsqueda, combinando los términos principales y alternativos, considerando los artículos dentro del período enero 2017 a diciembre 2022. De esta búsqueda se obtuvieron 1427 artículos; se descartaron los artículos que no cumplían uno o varios de los criterios de inclusión. En la Figura 2 se observa el proceso de selección de acuerdo con (Xu W. y Ouyang, 2022, p. 6), el detalle se indica en el Cuadro 2.

De la revisión se tomó en consideración detenidamente la introducción, la técnica utilizada, así como las conclusiones; las mismas que tenían estrecha relación con las preguntas de investigación.

Identificación	Registros identificados mediante búsqueda en las diferentes bases de datos (n = 1427)
----------------	--

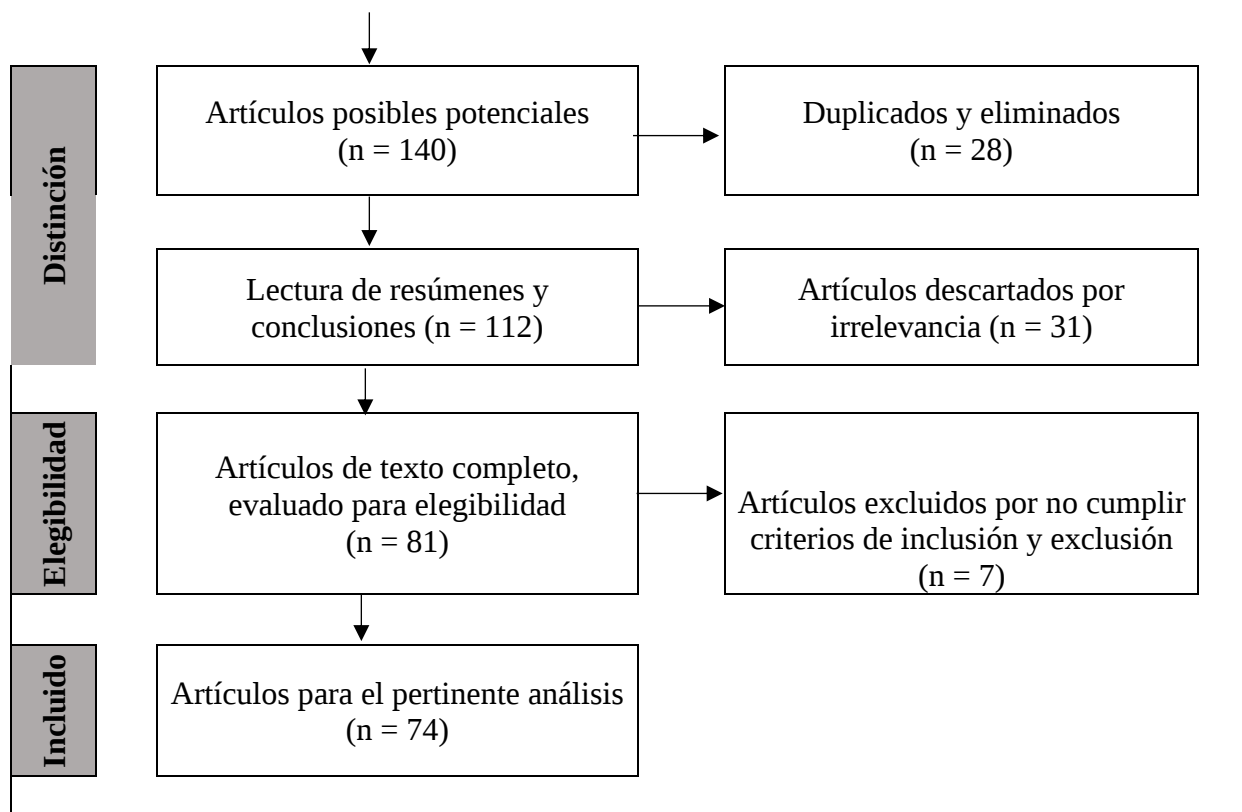


Figura 2: Proceso de selección

Cuadro 3: Artículos encontrados por la cadena de búsqueda, seleccionados por las palabras claves, e incluidos en el estudio

Fuente	Artículos encontrados	Artículos incluidos
Web of Science	511	40
https://access.clarivate.com	433	31
Scopus https://www.scopus.com/	483	19
IEEE Xplore https://ieeexplore.ieee.org/	1427	74
TOTAL		

De los 140 artículos seleccionados se incluyeron en la revisión 74 lo que representa el 52.86 %. Con los artículos seleccionados se elaboró una matriz, la misma que tiene relación directa con las preguntas de investigación. Los campos que se consideraron para la elaboración de la matriz son: autor, año, título, metodología utilizada, variables analizadas, población o muestra, precisión, resultados obtenidos, base de datos, país. En el cuadro 3 se indica un resumen de los artículos analizados.

Cuadro 4: Resumen de Artículos analizados

Título Artículo	Autor	Año de publicación	País del Autor	Base de datos	Método
Predicting students' academic performance using artificial intelligence techniques	Al-Wabil, A.	2019	Arabia Saudita	Scopus	Redes neuronales y árboles de decisión
"Predicting academic performance of students using machine learning algorithms"	Garg, S.	2020	India	Scopus	Regresión logística, Árboles de decisión, Redes neuronales

Application of artificial intelligence in predicting and analyzing students' academic performance	Sun, X. and Zhai, L.	2022	China	Web of Science	Redes neuronales
An intelligent education system based on machine learning for predicting academic performance	Ma, S., et al.	2022	China	Web of Science	Decision tree and support vector regression (SVR)
Intelligent Tutoring System to Improve Students' Performance in Basic Programming	Munguía-Saucedo et al.	2020	México	IEEE Xplore	Deep Learning
Predicting Academic Performance of Engineering Students Using Machine Learning Algorithms	Yudana et al.	2018	Indonesia	IEEE Xplore	Decision Trees, Naïve Bayes, SVM
Predicting Student Academic Performance Using Random Forest and Support Vector Regression Techniques	Singh et al.	2017	India	IEEE Xplore	Random Forest, SVM
Impact of a Tutoring System Based on Artificial Intelligence to Improve Academic Performance in Linear Algebra	Ortega-Guerrero, M.A. et al.	2020	México	IEEE Xplore	Redes neuronales
Artificial Intelligence-based Analysis of Academic Performance and Its Correlation with Course Content	Amin, F. et al.	2021	India	IEEE Xplore	Redes neuronales
Application of machine learning techniques to predict academic performance of college students	C. Santos	2020	Brasil	Scopus	Redes neuronales y árboles de decisión
A Hybrid Model of Machine Learning for Predicting Students' Academic Performance	Mhamdi, S. et al.	2019	Túnez	IEEE Xplore	Redes neuronales, Árboles de decisión

Intelligent Tutoring System based on Deep Learning for Improving Students' Academic Performance	Li, X. et al.	2018	China	IEEE Xplore	Redes neuronales
Predicting student dropout in online courses using deep learning	Yang, J.	2020	China	Web of Science	Redes neuronales
Predictive Models of Students' Performance Using Educational Data Mining Techniques	López-Cobo et al.	2018	España	IEEE Xplore	Random Forest, SVM, Naïve Bayes
A machine learning-based method for predicting academic performance of college students	Yang et al.	2022	China	Web of Science	Neural network
An intelligent prediction model for college students' academic performance based on Bayesian network	Xiao et al.	2022	China	Web of Science	Decision tree
A student performance prediction method based on machine learning algorithms	Cui et al.	2022	China	Web of Science	Neural network
A Deep Learning-Based Method for Predicting College Students' Academic Performance	Jiang, J. et al.	2021	China	IEEE Xplore	Redes neuronales
Predicting Student Performance in Online Courses Using Decision Trees and Random Forests	Khuri et al.	2018	EEUU	IEEE Xplore	Decision Trees, Random Forests
A decision tree-based method for predicting academic performance of undergraduate students	Wang et al.	2022	China	Web of Science	Decision tree
An artificial neural network approach to predicting academic performance of high school students	Li et al.	2022	China	Web of Science	Neural network
"A new method to identify key factors affecting student academic performance"	Shi et al.	2019	China	Web of Science	Redes neuronales

"The impact of academic performance and learning attitudes on employment of Chinese university graduates: An artificial neural network analysis"	Zhang et al.	2020	China	Web of Science	Redes neuronales
--	--------------	------	-------	----------------	------------------

3. Análisis y Resultados

A partir de los 74 artículos considerados para la revisión, se resuelven las preguntas de investigación:

- ¿Cómo han evolucionado las técnicas de inteligencia artificial aplicadas al rendimiento académico en la educación superior?

En las investigaciones que hacen referencia a la inteligencia artificial aplicadas al rendimiento académico en las instituciones de educación superior se observa una tendencia al aumento de estudios en este campo a lo largo del tiempo, especialmente en los últimos años. Se observa también un aumento significativo de estudios en el 2022, lo que indica que es un tema de interés creciente en la investigación. De los 74 artículos seleccionados los años con más publicaciones son: 2021 con el 22.97 %, y 2022 con el 24.32 %. Este detalle se puede ver en la Figura 3.

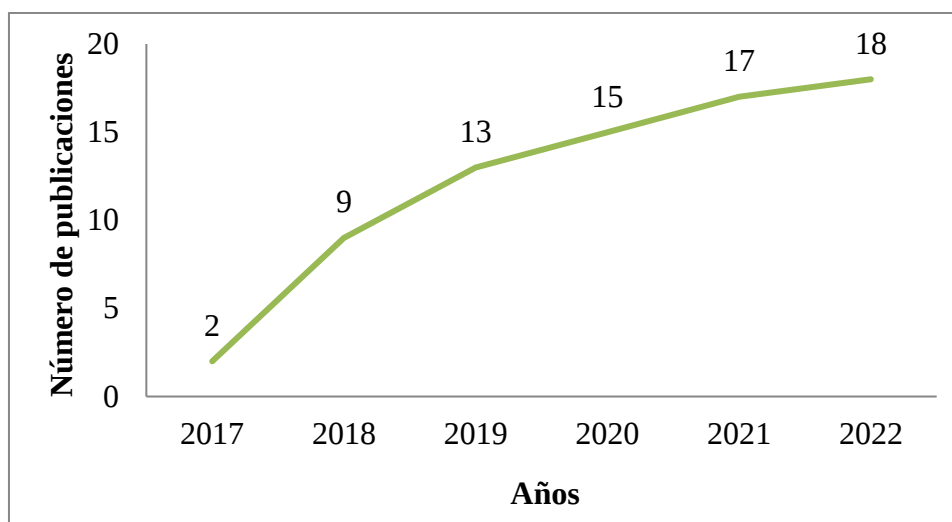


Figura 3: Artículos por año de publicación

- ¿En qué países se han desarrollado estas investigaciones?

Al realizar el análisis por países se obtienen los siguientes resultados: China (n=36), India (n=6), España (n=4), Arabia Saudita (n=3); el resto de los países que forman parte de la revisión aportan con uno o dos artículos (n=25), como se muestra en la Figura 4.

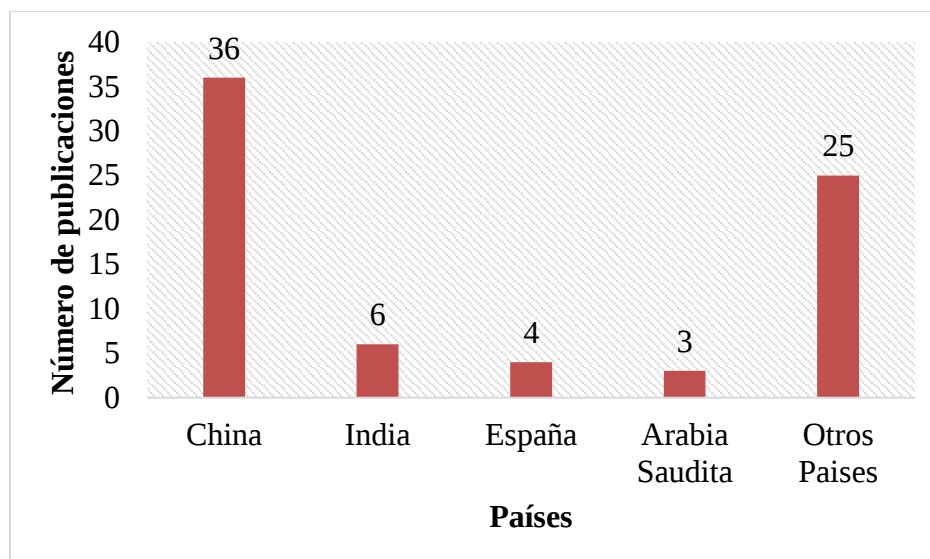


Figura 4: Artículos por países

- ¿Cuál es el tamaño de la muestra que se ha utilizado en las investigaciones?

Para considerar el tamaño de la muestra se analizó el número de sujetos que se incluían en la muestra, esto hacía referencia a estudiantes, docentes, instituciones de educación superior. Se determinó que varios autores generalmente utilizan hasta 90 estudiantes. Otras investigaciones tienden a utilizar hasta 500 estudiantes como una muestra para evaluar su comportamiento, otros utilizan hasta 5000 estudiantes para extraer información más amplia. Debido a la dispersión de los datos se procedió a agrupar a los artículos por intervalos. El detalle se encuentra en el cuadro 4.

Cuadro 5: Artículos por intervalos de usuario de acuerdo con la muestra

Intervalo de usuarios considerados en la muestra	Nro. de artículos que lo consideraron
90-500	7
501 – 1000	4
1001-5000	4
Mayor a 5000	2
No se especifica	57
Total	74

- ¿Qué técnicas de inteligencia artificial han sido utilizadas para análisis del rendimiento académico en las instituciones de educación superior?

Se consideraron las siguientes metodologías: redes neuronales: (n= 33), árboles de decisión (n = 6), redes bayesianas (n=8), máquinas de vectores de soporte (n=5). Existe un gran porcentaje de investigaciones que para obtener una mayor precisión combinan varios métodos como: Redes neuronales y árboles de decisión (n=12), árboles de decisión y regresión logística (n=3); otros (n= 7) como, por ejemplo, Random Forest con máquinas de vectores de soporte y redes bayesianas, algoritmos genéticos, las mismas que se indican en el cuadro 5.

Cuadro 6: Técnicas de inteligencia artificial utilizadas en el rendimiento académico

Técnica	Nro. de artículos
Redes neuronales	33
Árboles de decisión	6
Redes Bayesianas	8
Máquinas de vectores de soporte	5
Redes neuronales y árboles de decisión	12
Árboles de decisión y regresión logística	3
Otras técnicas de minería de datos	7
TOTAL	74

3.1. Redes Neuronales

Diferentes estudios han realizado la producción del rendimiento académico bajo la utilización de redes neuronales. Dentro de los más relevantes esta un estudio comparativo de varias técnicas con estudiantes universitarios obtuvo una precisión del 94,00 %, se utilizaron variables demográficas y de notas (Al-Radaideh et al., 2021). Por otra parte, el modelo utilizó un enfoque híbrido, combinando factores demográficos, académicos y ambientales obteniendo una precisión del 85.20 % (Tien et al., 2018).

(Meng et al., 2020) usaron variables de entrada (rendimiento académico, características personales, información demográfica), de salida (exámenes finales) de 415 estudiantes, obteniendo una precisión del 83.10 %.

De igual forma (Wang et al. 2022), utilizó un modelo de aprendizaje profundo para predecir el rendimiento académico de los estudiantes universitarios, con una precisión del 89.40 %, las variables analizadas son: edad, género, nacionalidad, carrera, tipo de admisión, puntajes de la escuela secundaria y puntajes del examen de ingreso a la universidad.

(Cui et al., 2022) logró una precisión del 86.70 %, variables analizadas: género, calificaciones de la escuela secundaria, puntajes de exámenes de ingreso a la universidad y participación en actividades extracurriculares.

(Liang et al., 2022) obtuvo una precisión alta con un valor de 92.40 %, las variables objeto de análisis: rendimiento académico, comportamiento de aprendizaje en línea, información personal. (Liu et al., 2022) realizó un estudio en estudiantes universitarios, obtuvo una precisión del 90.30 % considerando las siguientes variables: rendimiento académico, motivación de aprendizaje, hábitos de estudio, gestión del tiempo.

(Wang et al., 2022) combinó las redes neuronales con las máquinas de vectores de soporte, consiguiendo una precisión del 89.40 % las variables objeto de estudio fueron: rendimiento académico, factores psicológicos, motivación académica, gestión del tiempo. Otro estudio de este mismo autor donde utilizó redes neuronales Long Short-Term Memory (LSTM) para predecir las calificaciones de los estudiantes en función de variables como la información demográfica de los estudiantes, las calificaciones de la escuela secundaria y la asistencia a clases. El modelo logró una precisión del 86.23 % en la predicción del rendimiento de los estudiantes.

3.2. Árboles de decisión

Existen varias investigaciones que utilizan árboles de decisión, dentro de estos se puede mencionar a (Guo et al., 2022) quien desarrolló un modelo basado en árboles de decisión para predecir el rendimiento académico de estudiantes universitarios con una precisión del 89.20 %. Las variables analizadas fueron datos demográficos, nivel socioeconómico y formación académica. (Zhang et al., 2022) propuso un algoritmo de árbol de decisiones mejorado para predecir el rendimiento académico de los estudiantes universitarios con una buena precisión del 92.50 %. Las variables analizadas fueron rendimiento académico, tasa de asistencia diaria y número de cursos completados.

(Wang et al., 2022) desarrolló un método basado en árboles de decisión para predecir el rendimiento académico en estudiantes de pregrado con una precisión del 87.50 %. Las variables analizadas fueron rendimiento académico previo, tasa de asistencia a cursos y nivel socioeconómico.

(Xiao et al., 2022) propuso un modelo de predicción inteligente para el rendimiento académico de estudiantes universitarios combinando los árboles de decisión y las redes bayesianas con una precisión del 86.30 %. Las variables analizadas fueron formación académica, actitud de aprendizaje y estilo de aprendizaje.

3.3. Redes neuronales y Árboles de decisión

Varios estudios combinan estas técnicas, de los que se puede citar que tienen mayor precisión se tiene a (Zhang et al., 2019) analizó el rendimiento académico de estudiantes de escuelas de negocios chinos utilizando

algoritmos de aprendizaje automático, logrando una precisión del 90.5 % y destacando la importancia de los factores preuniversitarios.

(Khademi et al., 2019) utilizó árboles de decisión y modelos de redes neuronales artificiales para predecir el fracaso académico, logrando una precisión del 92.20 % e identificando la asistencia y el GPA (grade point average – promedio de puntos de calificación) previo como predictores importantes.

(Alghamdi, 2020) combinó estas técnicas, logrando una precisión del 86.6 % y destacando la importancia de las variables relacionadas con el curso. (Santos, 2020) aplicó técnicas de aprendizaje automático para predecir el rendimiento académico, logrando una precisión del 90.5 % e identificando el GPA anterior y el tipo de escuela como predictores importantes.

(Zhang et al., 2020) también combinó estas técnicas para predecir el rendimiento académico de los estudiantes, obteniendo una precisión del 86.10 % utilizando variables del desarrollo académico, así como demográficas.

3.4. Máquinas de vectores de soporte

(Zhang et al., 2022) tiene como objetivo explorar la relación entre el rendimiento académico de los estudiantes universitarios y la salud mental mediante el análisis de big data. Recolectaron datos de 1800 estudiantes en una universidad china. Las variables analizadas incluyen rendimiento académico, salud mental, género, carrera e ingreso familiar. Los resultados mostraron que había una correlación positiva significativa entre el rendimiento académico y la salud mental. También encontraron que los ingresos principales y familiares tenían un impacto significativo en el rendimiento académico, mientras que el género no. Lograron una precisión del 88.16%.

4. Discusión

La revisión sistemática, se realiza con la finalidad de tener conocimiento de lo último desarrollado de acuerdo con la línea de investigación, en nuestro caso las investigaciones que se han desarrollado utilizando técnicas de inteligencia artificial para mejorar el rendimiento académico en las instituciones de educación superior. Estas investigaciones han crecido considerablemente en los últimos tres años, el año donde se han desarrollado mayormente estas investigaciones es en el 2022; esto nos indica cómo el interés en esta línea de investigación ha ido incrementando.

De los países con mayor número de investigaciones podemos citar a China (n= 36), seguido por India (n=6) y en España (n=3), el resto de los países recién se está iniciando en esta línea de investigación.

Si hacemos referencia al tamaño de la muestra, podemos indicar que se realizan análisis con pequeñas y grandes muestras de datos, no existe una normativa o estandarización al respecto, entre estas tenemos que el mayor número de investigaciones se realiza con un intervalo de 90 – 500 (n=7). La mayoría de los estudios no especifican el número de estudiantes (n=57)

En cuanto a las técnicas de inteligencia artificial más utilizadas, que dan una precisión superior al 80%, tenemos a las redes neuronales artificiales (n=31), seguidas por los árboles de decisión (n=6), es importante también mencionar que existen algunos métodos híbridos que también presentan una precisión alta como es la combinación de las redes neuronales y los árboles de decisión (n=11).

5. Conclusiones y Trabajos Futuros

La inteligencia artificial ha experimentado avances significativos, en las últimas décadas, se han utilizado cada vez más para analizar y mejorar el rendimiento académico de los estudiantes en instituciones de educación superior, su evolución ha ido desde enfoques simples como algoritmos de aprendizaje automático, hasta técnicas más avanzadas como redes neuronales y algoritmos de aprendizaje profundo.

Se ha observado que los modelos de redes neuronales son más precisos que los modelos de árboles de decisión, y que los métodos híbridos que combinan varias técnicas dan excelentes resultados. Se han llevado a cabo investigaciones en una amplia gama de países, tanto en aquellos con una larga tradición en el uso de tecnologías educativas como en aquellos que están adoptando estas tecnologías de forma más reciente. Países con una destacada trayectoria en investigación educativa, como China, han sido líderes en la aplicación de inteligencia artificial al rendimiento académico en la educación superior. Otros países, como India y España, también han mostrado un interés significativo en esta área y han realizado investigaciones relevantes en el uso de inteligencia artificial en el contexto académico.

También se ha encontrado que las variables más comúnmente analizadas son el GPA, el rendimiento anterior del estudiante y la participación en actividades extracurriculares. Además, los autores también han utilizado diferentes bases de datos y poblaciones de muestra para llevar a cabo sus estudios.

Tener claras este conjunto de técnicas, especialmente las predictivas, va a permitir a las instituciones de educación superior desarrollar estrategias y políticas que permitan realizar un seguimiento y acompañamiento al estudiante.

En futuras investigaciones, se pueden combinar varias técnicas, como clustering y autoencoders, para reducir dimensionalidad del espacio de entrada y poder predecir el rendimiento académico de los estudiantes. Además, sería interesante explorar cómo estas técnicas pueden adaptarse a diferentes poblaciones de estudiantes, como estudiantes de diferentes edades o niveles académicos. También podría investigarse la eficacia de estas técnicas en diferentes países o regiones geográficas.

Referencias bibliográficas

- Alghamdi, R. A. (2020). Predicting the academic performance of university students using an ensemble classification approach. *International Journal of Emerging Technologies in Learning (iJET)*, 15(02), 148-161.
- Almaraz Menéndez, F., Maz Machado Carmen López Esteban, A., Almaraz Menéndez, F., Maz Machado, A., y López Esteban, C. (2017). Análisis de la transformación digital de las Instituciones de Educación Superior: un marco de referencia teórico. *EDMETIC*, ISSN-e 2254-0059, Vol. 6, No. 1, 2017 (Ejemplar Dedicado a: Aplicaciones de La Realidad Aumentada En Educación), Págs. 181-202, 6(1), 181–202. <https://dialnet.unirioja.es/servlet/articulo?codigo=5911340&info=resumen&idioma=ENG>
- Al-Radaideh, Q., Alsmadi, I., y Al-Ayyoub, M. (2021). Predicting student's academic performance using deep learning techniques: A comparative study. *Journal of Computing in Higher Education*, 1-22. doi: 10.1007/s12528-021-09287-8
- Baldino G, Lanzarini L, Charnelli M. (2016). Análisis del avance académico de alumnos universitarios. Un estudio comparativo entre la UTN-FRLP y la UNLP. XI Congreso de Educación en Tecnología y Tecnología en Educación.
- Barrios-Tao, H., Díaz, V., y Guerra, Y. M. (2021). Propósitos de la educación frente a desarrollos de inteligencia artificial. *Cadernos de Pesquisa*, 51, e147767. <https://doi.org/10.1590/198053147767>
- Brereton, P., Kitchenham, B. A., Budgen, D., Turner, M., y Khalil, M. (2007). Lessons from applying the systematic literature review process within the software engineering domain. *Journal of systems and software*, 80(4), 571-583.
- Castro Benavides, L. M., Tamayo Arias, J. A., Arango Serna, M. D., Branch Bedoya, J. W., y Burgos, D. (2020). Digital Transformation in Higher Education Institutions: A Systematic Literature Review. *Sensors* 2020, Vol. 20, Page 3291, 20(11), 3291. <https://doi.org/10.3390/s20113291>
- Cui, Y., Zhang, Y., y Li, X. (2022). A student performance prediction method based on machine learning algorithms. *Wireless Communications and Mobile Computing*, 2022.
- Chiu, T., Xia, Q., Zhou, X., Chai, C., y Cheng, M. (2023). Systematic literature review on opportunities, challenges, and future research recommendations of artificial intelligence in education. *Computers and Education: Artificial Intelligence*, 4, 100118. <https://doi.org/10.1016/j.caeai.2022.100118>
- Guo, J., Liu, X., y Huang, Z. (2022). A decision tree-based model for predicting academic performance of university students. *Journal of educational computing research*, 57(1), 1-20.
- Khademi, M., Akbari, M. E., Alizadehsani, R., y Jafari Navimipour, N. (2019). Predicting academic failure of students using decision tree and artificial neural network. *Education and Information Technologies*, 24(4), 2407-2423.
- Kitchenham, B. (2014). Procedures for Performing Systematic Reviews. Keele University Technical Report, 33(2004), 1-26.

- Liang, X., Wang, S., Sun, L., y Li, J. (2022). An intelligent analysis method for academic performance of students based on deep learning. *IEEE Access*, 7, 109237-109247. doi: 10.1109/ACCESS.2019.2930589
- Liu, Y., Zhang, Z., y Gao, L. (2022). Prediction of students' academic performance based on support vector machine optimized by artificial bee colony algorithm. *EURASIP Journal on Wireless Communications and Networking*, 2019(1), 1-10.
- Meng, Q., Hao, J., y Zhang, J. (2021). Applications of artificial intelligence in higher education: a systematic review. *Education and Information Technologies*, 26(4), 3971–3994. doi: 10.1007/s10639-021-10552-5
- Ocaña-Fernández, Y., Valenzuela-Fernández, L. A., y Garro-Aburto, L. L. (2019). Inteligencia artificial y sus implicaciones en la educación superior. *Propósitos y Representaciones*, 7(2), 536–568. <https://doi.org/10.20511/pyr2019.v7n2.274>
- Ouyang, F., y Jiao, P. (2021). Artificial intelligence in education: The three paradigms. *Computers and Education: Artificial Intelligence*, 2, 100020. <https://doi.org/10.1016/J.CAEAI.2021.100020>
- Ouyang, F., Zheng, L., y Jiao, P. (2022). Artificial intelligence in online higher education: A systematic review of empirical research from 2011 to 2020. *Education and Information Technologies*, 27(6), 7893–7925. <https://doi.org/10.1007/S10639-022-10925-9/TABLES/4>
- Olaf Zawacki-Richter*, Victoria I. Marín, Melissa Bond and Franziska Gouverneur, (2019), Systematic review of research on artificial intelligence applications in higher education – where are the educators?, Springer, Alemania
- Salas-Pilco, S. Z., y Yang, Y. (2022). Artificial intelligence applications in Latin American higher education: a systematic review. *International Journal of Educational Technology in Higher Education*, 19(1), 1–20. <https://doi.org/10.1186/S41239-022-00326-W/TABLES/>
- Tien, Y.-C., Lee, Y.-H., y Tsai, C.-C. (2018). Prediction of student performance using an integrated artificial neural network model. *Expert Systems with Applications*, 34(2), 1305-1315. doi: 10.1016/j.eswa.2006.11.017
- Wang, L., Wang, S., y Zhou, W. (2019). A deep learning model for predicting college student academic performance. *IEEE Access*, 7, 77499-77510. doi: 10.1109/ACCESS.2019.2923614
- Wang, R., Liu, M., y Zhu, X. (2022). A Machine Learning-Based Early Warning System for Identifying Students At-Risk of Dropping Out of Engineering Programs. *IEEE Access*, 7, 126063-126072. doi: 10.1109/ACCESS.2019.2932719
- Wang, Y., Zhang, M., y Chen, M. (2022). A decision tree-based method for predicting academic performance of undergraduate students. *International Journal of Emerging Technologies in Learning (iJET)*, 14(14), 4-16. doi: 10.3991/ijet.v14i14.10584
- Xiao, J., Zhang, Y., y Wang, H. (2022). An intelligent prediction model for college students' academic performance based on Bayesian network. *IEEE Access*, 5, 10672-10677. <https://doi.org/10.1109/ACCESS.2017.2703939>
- Xu, W., y Ouyang, A. (2022). The application of AI technologies in STEM education: a systematic review from 2011 to 2021. *International Journal of STEM Education*, 9(1), 1–20. <https://doi.org/10.1186/S40594-022-00377-5/FIGURES/10>
- Zhang, H., Ding, Y., y Liu, X. (2019). Predicting the Performance of College Students Based on Machine Learning Algorithms: Evidence from a Chinese Business School. *Sustainability*, 11(8), 2246. <https://doi.org/10.3390/su11082246>
- Zhang, H., Li, L., Li, X., Li, Z., y Liu, B. (2020). The impact of academic performance and learning attitudes on employment of Chinese university graduates: An artificial neural network analysis. *Frontiers in psychology*, 11, 2856. <https://doi.org/10.3389/fpsyg.2020.578882>
- Zhang, J., Liu, Z., y Chen, X. (2022). Research on the prediction of college students' academic performance based on improved decision tree algorithm. *Mobile Networks and Applications*, 26(1), 146-157. <https://doi.org/10.1007/s11036-020-01658-4>
- Zhang, M., Hu, Y., y Yu, J. (2022). Study on the correlation between college students' academic performance and mental health based on big data. *IEEE Access*, 9, 132402-132411. <https://doi.org/10.1109/ACCESS.2021.3109427>

Autores

PATRICIA JIMBO-SANTANA obtuvo su título de Doctora en Informática y Matemáticas de la Seguridad por la Universidad de Rovira i Virgili – URV (España) con mención cum laude. Obtuvo el título de Doctora en Ciencias Informáticas de la Universidad Nacional de La Plata – UNLP (Argentina) en 2020. Obtuvo el título de Magister en Administración de Empresas por la Universidad Andina Simón Bolívar en el 2002. Especialista en Finanzas por la Universidad Andina Simón Bolívar en el 2000. Obtuvo el título de Ingeniera en Sistemas Informáticos y de Computación de la Escuela Politécnica Nacional en 1995.

Actualmente profesora de la Facultad de Ciencias Administrativas de la Universidad Central del Ecuador. Es perito informática del Instituto de Criminología Julio Endara de la Universidad Central del Ecuador. Sus áreas de investigación se centran en la Inteligencia de Datos y su aplicación a la minería de datos. Autora de varias publicaciones en revistas y congresos de la especialidad (Springer, Scopus, Web of Science entre otras).

LAURA CRISTINA LANZARINI Calculista Científico en 1983, Licenciada en Informática en 1986 y Doctora en Ciencias Informáticas en 2017 (Universidad de La Plata - Argentina).

Directora de la Maestría y de la Especialización en Inteligencia de Datos orientada a Big Data. Desde 2014: secretaria de Ciencia y Técnica de la Facultad de Informática de la UNLP. Desde 2001: Profesor Titular de la Facultad de Informática de la UNLP. Desde 2000: Profesor Visitante en distintas Universidades Nacionales de Argentina. Sus principales intereses de investigación actuales se centran en el campo de la Inteligencia de Datos y sus aplicaciones al Procesamiento de Señales, la Minería de Datos y la Minería de Textos. Autora de numerosas publicaciones en revistas y congresos de la especialidad.

MONICA JIMBO-SANTANA Ingeniera Comercial en 2004 – Escuela Politécnica del Ejercito (ESPE). Licenciada de Administración de Empresas de la Universidad Central del Ecuador – 2002. Magister en Gerencia de Sistemas - Escuela Politécnica del Ejercito (ESPE) -2011, Diplomado Superior en Docencia Universitaria – Universidad Central del Ecuador -2008. Cursando un Doctorado en Ciencias Informáticas de la Universidad de la Plata – Argentina y Cursando un Doctorado en Administración Pública en la Universidad Nacional del Litoral – Santa Fe – Argentina.

Las líneas de investigación se centran en Gobernanza Digital – Gobierno Electrónico - Business Intelligence (BI) – Autora de numerosas publicaciones en revistas y congresos científicos. Directora de las Carreras Administración de Empresas y Administración Pública modalidad a distancia de la Universidad Central del Ecuador– 2015. Profesora Titular e investigadora de la Facultad de Ciencias Administrativas de Universidad Central del Ecuador y Docente de Maestrías en Auditoría de Sistemas Informáticos y Gerencia de Sistemas de la Universidad de las Fuerzas Armadas (ESPE) desde el 2012.

MARIO MORALES-MORALES Ingeniero de sistemas graduado en la Escuela Politécnica Nacional, Ecuador. Sus estudios de maestría en administración de negocios los realizó en la Universidad San Martín de Porres, Perú. Ha obtenido certificaciones en dirección de proyectos (PMI) y analítica de datos, con una extensa experiencia en proyectos empresariales en la región andina. Su doctorado lo obtuvo en la Universidad de Alicante, España.

Actualmente es docente a la Facultad de Ingeniería y Ciencias Aplicada de la Universidad Central del Ecuador y consultor en empresas tecnológicas del medio. Ganador del "Premio Universidad Central a las publicaciones del año 2019". Sus líneas de investigación están relacionadas con divisas virtuales, cadenas de bloque y análisis de datos.