



REVISTA

CÁTEDRA

Artificial intelligence for analyzing academic performance in higher education institutions. A systematic literature review

*Inteligencia artificial para analizar el rendimiento académico en instituciones de educación superior.
Una revisión sistemática de la literatura*

Patricia Jimbo-Santana

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

prjimbo@uce.edu.ec

<https://orcid.org/0000-0001-7432-1622>

Laura Cristina Lanzarini

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

laural@lidi.info.unlp.edu.ar

<https://orcid.org/0000-0001-7027-7564>

Mónica Jimbo-Santana

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

djimbo@uce.edu.ec

<https://orcid.org/0000-0002-3948-9507>

Mario Morales-Morales

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

mmoralesm@uce.edu.ec

<https://orcid.org/0000-0002-7493-8072>

(Received on: 17/04/2023; Accepted on: 26/05/2023; Final version received on: 15/07/2023)



[Licencia Creative Commons Atribución 4.0 Internacional \(CC BY 4.0\)](https://creativecommons.org/licenses/by/4.0/)

Suggested citation: Jimbo-Santana, P., Lanzarine, L.C., Jimbo-Santana, M. y Morales-Mroales, M. (2023). Artificial intelligence for analyzing academic performance in higher education institutions. A systematic literature review. *Revista Cátedra*, 6(2), 30-51.

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.

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.



[Licencia Creative Commons Atribución 4.0 Internacional \(CC BY 4.0\)](https://creativecommons.org/licenses/by/4.0/)

1. Introduction

Today we are immersed in a society that is increasingly subject to large-scale technological processes:

The area of education (which is sensitive to changes in society as it advances along with it) is also going through such a trend of adaptation to the novel communities of technological interaction; a process that is oriented to new trends and profiles in relation to the new proposals in the sector (Ocaña, 2019, p. 537).

Artificial intelligence represents a synonym of innovation and technology, since it mimics the processes of human intelligence with the use of algorithms and helps to find patterns of behavior that are not visible to the naked eye. In the field of education it has been of great help in various fields, including the prediction of academic performance, thanks to the fact that artificial intelligence is responsible for processing large amounts of data, which allows us to better understand the profiles of students and thus their needs. Based on these requirements, plans, projects and academic activities can be established to help improve the quality of education, understand the profile of students and their needs in terms of the level of progress or beginnings of desertion, which is important because in this way the institutions can make decisions that provide support in the learning processes, and at the same time maintain a constant monitoring of the level of quality of education.

If we refer to educational theories, we can mention the instructional design theory that uses data collected through digital technology and artificial intelligence to inform the design of effective learning environments. It has been observed that artificial intelligence is able to analyze data on students' academic performance and provide valuable information to improve the design of materials, activities and assessments. Another relevant theory is that of adaptive learning, which proposes adapting instruction and educational resources according to the individual needs of each student. Using machine learning algorithms, artificial intelligence can analyze data on academic performance and adjust content delivery, activity sequencing, and teaching strategies to optimize learning for each student.

Artificial intelligence-based applications for education have been growing rapidly around the world, and have impacted higher education institutions. Several of these applications are being implemented to improve university services, help teachers provide quality education, make decisions regarding university attrition, and support student learning (Salas-Pilco and Yang, 2022).

Higher education institutions have detailed information on each student, however, they lack models that allow them to describe the student objectively, and based on this analysis make decisions that allow them to avoid problems such as student dropout (Baldino et al., 2016, p. 589).

On the other hand, over time it has been possible to evidence a steady growth of artificial intelligence applications in educational processes, being one of the emerging fields educational technology, because although it has existed for approximately 30 years, it is still unclear whether it can significantly impact teaching and learning in higher education (Zawacki-Richter et al., 2019, p. 2).

Some of the problems that may arise in higher education institutions by not using artificial intelligence in the assessment of academic performance may include: lack of efficiency in assessment, difficulty in identifying patterns and trends, lack of early intervention in



[Licencia Creative Commons Atribución 4.0 Internacional \(CC BY 4.0\)](https://creativecommons.org/licenses/by/4.0/)

student performance, limitations in institutional decision making, among others. It should not be overlooked that the education system is shifting from focusing on product to process, expanding beyond the domain of knowledge to include self-regulation, collaboration, and motivation (Roll and Wylie, 2016).

In research using artificial intelligence to predict academic performance, several variables are used to build prediction models, such as the student's academic history, age, gender, socioeconomic status, among others. These variables are processed by machine learning algorithms that seek to identify patterns and relationships between them, and thus predict the student's future academic performance.

Currently, not employing artificial intelligence techniques in academic performance could imply a lack of efficiency in information management, create biases and subjectivities, as well as lack of personalization and adaptation to the specific needs of students.

The objective of the present work is to perform a systematic literature review that considers research conducted from January 2017 to December 2022, on artificial intelligence techniques (data mining, deep learning, big data), to analyze academic performance in university institutions, considered as the result obtained by a student in one or several standardized tests, such as exams, homework, projects, among others, which measure the level of knowledge and skills acquired in a certain period of time. Academic performance can also be measured in terms of class attendance, participation in extracurricular activities, fulfillment of homework and school responsibilities, performance in teamwork, and other indicators that reflect the student's academic performance.

This paper is organized as follows: section 2 describes the materials and methods used, section 3 indicates the analysis performed and the results obtained, section 4 deals with the Discussion, and finally section 5 indicates the conclusions and future work.

2. Materials and Methods

It is important to first consider the objectives of this research:

- To analyze the evolution of artificial intelligence techniques applied to academic performance in higher education.
- To identify the countries in which these investigations have been developed.
- To evaluate the sample size used in research on the analysis of academic performance in higher education institutions.
- To identify and describe the artificial intelligence techniques used for the analysis of academic performance in higher education institutions.

2.1 Methods

The research used is a systematic review of the literature, which aims to synthesize and critically analyze the available evidence on a specific topic. This involves carrying out a set of activities framed in three phases, i) approach, ii) conduct, and, iii) presentation, following the scheme proposed by (Kitchenham, 2014) and that gives formal support to this research (Figure 1). To this end, a systematic and exhaustive search is carried out in databases and other relevant sources, predefined inclusion and exclusion criteria are applied and the selected studies are analyzed.



[Licencia Creative Commons Atribución 4.0 Internacional \(CC BY 4.0\)](https://creativecommons.org/licenses/by/4.0/)

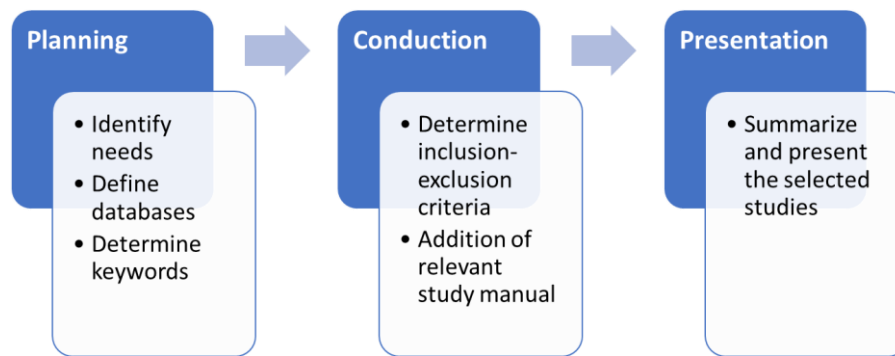


Figure 1. Systematic review activities

The techniques and strategies for the study included searching the scientific databases Web of Science, Scopus and IEEE Xplore, with keywords related to artificial intelligence and academic performance. Articles published from January 2017 to December 2022 were selected and predefined inclusion and exclusion criteria were applied to select the 74 articles analyzed. In addition, a critical and analytical approach was used to synthesize the results and discuss the implications and limitations of the studies found.

According to Brereton (2007) "conducting a systematic review involves several discrete activities, which can be grouped into three main phases: planning, conducting the review, and reporting the review" (p. 2), which is considered to determine the current status of artificial intelligence, in relation to academic performance in higher education.

In the execution of the protocol, the need for the review was identified, the questions that will guide the research were determined, the terms, strategies and search strings (English and Spanish) were identified, the inclusion criteria were established to proceed with the data extraction. Finally, the review protocol was executed and developed.

2.2 Necessity of the revision

In conducting the respective review, several systematic review articles were found on the use of artificial intelligence in higher education, focused on various fields, thus we can indicate:

Chiu (2023) performs a systematic review between 2012 - 2021 "emphasizing four domains: learning, teaching, evaluation and administration" (p. 1). In another study of conducted a systematic review between 2012 - 2020 where they focused on online learning, the same that has been implemented gradually in higher education in recent years, in addition to bringing challenges and new challenges, also allows institutions to improve instruction and learning (Ouyang et al., 2022, p. 23).

They also conducted a literature review, analyzing articles from 1980 to 2019, here they mention the technological advancement that brings the Industrial Revolution 4.0, this allows higher education institutions to be forced to face the digital transformation and consider within it artificial intelligence (Castro-Benavides et al., 2020, p.1). In another work of perform a theoretical framework, analyzing the digital transformation in higher education institutions (Almaraz Menéndez et al., 2017).

Other research focuses on the changes that have been generated in higher education institutions, this due to artificial intelligence, in addition to the new challenges that come hand in hand with it, thus concluding that artificial intelligence is a challenge of great importance for higher education institutions, but nevertheless in the case of Latin America



[Licencia Creative Commons Atribución 4.0 Internacional \(CC BY 4.0\)](https://creativecommons.org/licenses/by/4.0/)

this reality is still unattainable (Ocaña Fernández, et al., 2019, p. 8). Another research focuses on the opportunities, risks and possibilities faced by education, and the advantages obtained by using artificial intelligence (Barrios Tao et al., 2021, p. 1). Another study talks about the evolution that has been evidenced in technology, and how artificial intelligence has been applied in different fields in higher education institutions (Ouyang and Jiao, 2021).

Of these investigations, none of them provide significant information in the area of academic performance in higher education institutions. It is important to consider that technology is advancing day by day, so it is necessary to conduct a literature review that incorporates the latest research.

2.3 Research Questions

The research questions guiding the systematic review are:

- How have artificial intelligence techniques applied to academic performance in higher education evolved?
- In which countries has this research been developed?
- What is the sample size that has been used in the investigations?
- What artificial intelligence techniques have been used for the analysis of academic performance in higher education institutions?

2.4 Search terms

Based on a preliminary literature survey, the search terms contained in the title, keywords and abstract of the articles were identified through different combinations of main and secondary terms, as shown in Table 1, and AND and OR logical operators were used.

The English search string used was:

("artificial intelligence" OR AI) AND ("academic performance" OR "college performance") AND ("higher education" OR university OR college).

Main terms in Spanish	Alternative terms in Spanish	Main terms in English	Alternative terms in English
Artificial Intelligence	IA	Artificial Intelligence	AI
Academic Performance	University performance	Academic Performance	College Performance
Higher education	University	Higher Education	University College

Table 1. Search Terms

2.5 Search strategies

The selected databases included Web of Science, Scopus, and IEEE Xplore, which were selected for their relevance to this line of research.

2.6 Inclusion criteria

The articles were selected according to the following criteria:



[Licencia Creative Commons Atribución 4.0 Internacional \(CC BY 4.0\)](https://creativecommons.org/licenses/by/4.0/)

Inclusion criteria	Exclusion criteria
<ul style="list-style-type: none"> • Scientific articles should be related to the application of artificial intelligence to predict and/or describe the academic performance of students in higher education institutions. • Scientific articles should relate to the application of artificial intelligence to predict and/or describe the academic performance of students in higher education institutions. • Scientific studies or articles are relevant to answer the research questions. • Scientific articles should cover the entire university field, i.e., without distinction of country, university. Careers or professions. • Studies or scientific articles are relevant to answer the research questions. • Studies should be within the research range from January 2017 to December 2022. • Studies must be in English and Spanish language. • Articles that include search strings in the title, keywords, or abstract. • Articles published in journals or conferences • Scientific articles available in full text 	<ul style="list-style-type: none"> • Studies that applied artificial intelligence to predict or forecast academic performance in education in schools, colleges and kindergartens. • Scientific articles focusing on artificial intelligence tools for the teaching-learning process, engineering, physical education, programming, among others. • Studies conducted before 2017.

Table 2. Inclusion criterio

2.7 Data extraction

Several search strings were structured, combining main and alternative terms, considering articles within the period January 2017 to December 2022. From this search, 1427 articles



[Licencia Creative Commons Atribución 4.0 Internacional \(CC BY 4.0\)](https://creativecommons.org/licenses/by/4.0/)

were obtained; articles that did not meet one or more of the inclusion criteria were discarded. Figure 2 shows the selection process according to (Xu W. and Ouyang, 2022, p. 6), the detail is shown in Table 2.

From the review, careful consideration was given to the introduction, the technique used, as well as the conclusions, which were closely related to the research questions. Figure 2. Selection process

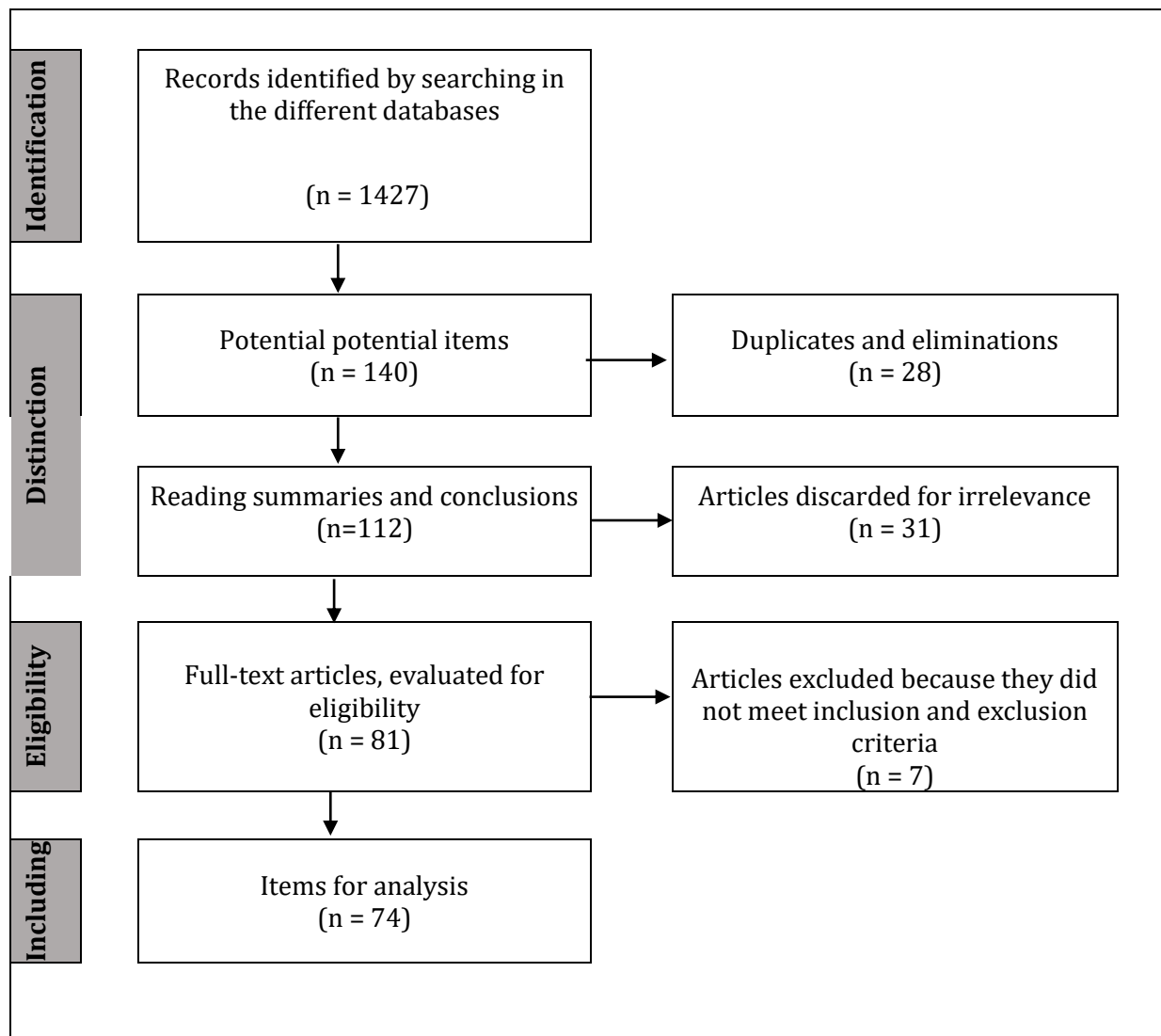


Figure 2. Selection process



[Licencia Creative Commons Atribución 4.0 Internacional \(CC BY 4.0\)](https://creativecommons.org/licenses/by/4.0/)

Source	Found articles	Items included
Web of Science https://access.clarivate.com	511	40
Scopus https://www.scopus.com/	433	31
IEEE Xplore https://ieeexplore.ieee.org/	483	19
TOTAL	1427	74

Table 3. Articles found by the search string, selected by keywords, and included in the study.

Of the 140 articles selected, 74 were included in the review, representing 52.86%. A matrix was prepared with the selected articles, which is directly related to the research questions. The fields considered for the elaboration of the matrix are: author, year, title, methodology used, variables analyzed, population or sample, precision, results obtained, database, country. A summary of the articles analyzed is shown in Table 3.

Title Article	Author	Year of publication	Author's country	Database	Method
Predicting students' academic performance using artificial intelligence techniques	Al-Wabil, A.	2019	Saudi Arabia	Scopus	Neural Networks and Decision Trees
"Predicting academic performance of students using machine learning algorithms"	Garg, S.	2020	India	Scopus	Logistic regression, Decision trees, Neural networks, Neural networks
Application of artificial intelligence in predicting and analyzing students' academic performance	Sun, X. and Zhai, L.	2022	China	Web of Science	
An intelligent education system based on machine learning for	Ma, S., et al.	2022	China	Web of Science	Decision tree and support vector



[Licencia Creative Commons Atribución 4.0 Internacional \(CC BY 4.0\)](https://creativecommons.org/licenses/by/4.0/)

Title Article	Author	Year of publication	Author's country	Database	Method
predicting academic performance					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	Neural Networks
Application of machine learning	C. Santos	2020	Brasil	Scopus	Neural Networks



[Licencia Creative Commons Atribución 4.0 Internacional \(CC BY 4.0\)](https://creativecommons.org/licenses/by/4.0/)

Title Article	Author	Year of publication	Author's country	Database	Method
techniques to predict academic performance of college students					and Decision Trees
A Hybrid Model of Machine Learning for Predicting Students' Academic Performance	Mhamdi, S. et al.	2019	Túnez	IEEE Xplore	Neural networks, Decision trees
Intelligent Tutoring System based on Deep Learning for Improving Students' Academic Performance	Li, X. et al.	2018	China	IEEE Xplore	Neural networks
Predicting student dropout in online courses using deep learning	Yang, J.	2020	China	Web of Science	Neural Networks
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



[Licencia Creative Commons Atribución 4.0 Internacional \(CC BY 4.0\)](https://creativecommons.org/licenses/by/4.0/)

Title Article	Author	Year of publication	Author's country	Database	Method
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	Neural networks
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	Neural Networks
"The impact of academic performance and learning attitudes on employment of Chinese university	Zhang et al.	2020	China	Web of Science	Neural Networks



[Licencia Creative Commons Atribución 4.0 Internacional \(CC BY 4.0\)](https://creativecommons.org/licenses/by/4.0/)

Title Article	Author	Year of publication	Author's country	Database	Method
graduates: An artificial neural network analysis"					

Table 4. Summary of analyzed articles



[Licencia Creative Commons Atribución 4.0 Internacional \(CC BY 4.0\)](https://creativecommons.org/licenses/by/4.0/)

3. Analysis and Results

Based on the 74 articles considered for the review, the following research questions are resolved:

- How have artificial intelligence techniques applied to academic performance in higher education evolved?

In the research that refers to artificial intelligence applied to academic performance in higher education institutions, there is a trend of increasing studies in this field over time, especially in recent years. A significant increase of studies is also observed in 2022, indicating that it is a topic of growing interest in research. Of the 74 articles selected the years with the most publications are: 2021 with 22.97 %, and 2022 with 24.32 %. This detail can be seen in Figure 3.

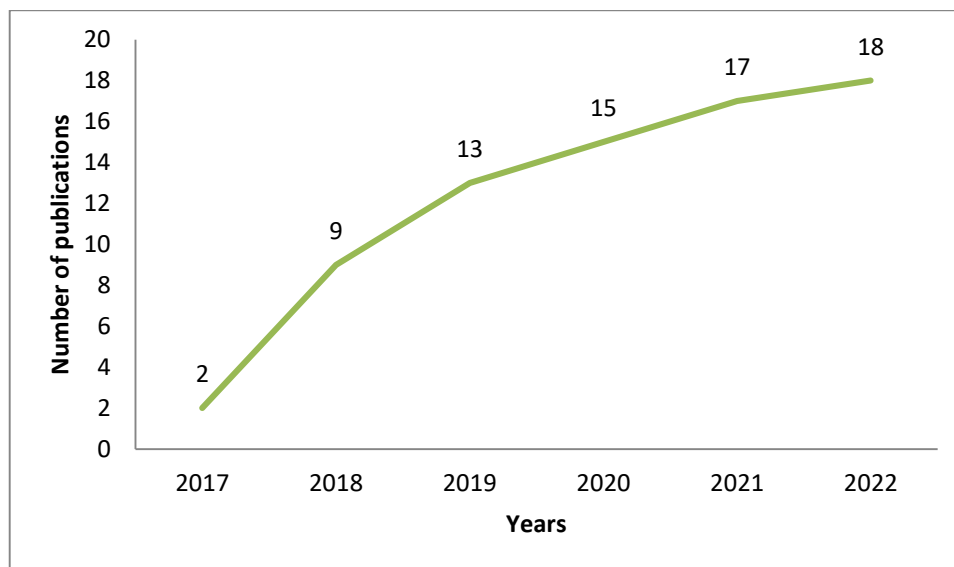


Figure 3. Articles by year of publication

- In which countries has this research been carried out?

The analysis by country shows the following results: China (n=36), India (n=6), Spain (n=4), Saudi Arabia (n=3); the rest of the countries included in the review contributed one or two articles (n=25), as shown in Figure 4.



[Licencia Creative Commons Atribución 4.0 Internacional \(CC BY 4.0\)](https://creativecommons.org/licenses/by/4.0/)

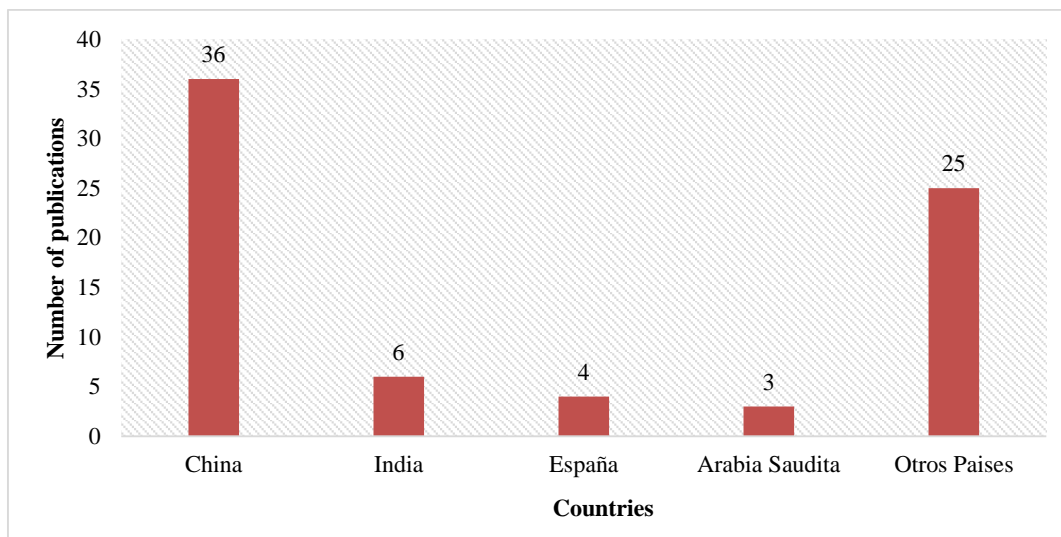


Figure 4. Articles by country

- What is the sample size used in the research?

To consider the sample size, the number of subjects included in the sample was analyzed, this referred to students, teachers, higher education institutions. It was determined that several authors generally use up to 90 students. Other research tends to use up to 500 students as a sample to evaluate their behavior, others use up to 5000 students to extract broader information. Due to the dispersion of the data, the articles were grouped by intervals. The details can be found in Table 4.

Range of users considered in the sample	No. of articles that considered it
90-500	7
501 – 1000	4
1001-5000	4
Greater than 5000	2
Not specified	57
Total	74

Table 5. Items by user intervals according to the sample.

- What artificial intelligence techniques have been used for academic performance analysis in higher education institutions?

The following methodologies were considered: neural networks (n=33), decision trees (n=6), Bayesian networks (n=8), support vector machines (n=5). There is a large percentage of studies that combine several methods to obtain greater precision, such as: neural networks and decision trees (n=12), decision trees and logistic regression (n=3); others (n=7), such as Random Forest with support vector machines and Bayesian networks, genetic algorithms, as shown in Table 5.



[Licencia Creative Commons Atribución 4.0 Internacional \(CC BY 4.0\)](https://creativecommons.org/licenses/by/4.0/)

Technique	No. of items
Neural Networks	33
Decision trees	6
Bayesian networks	8
Support Vector Machines	5
Neural networks and decision trees	12
Decision trees and logistic regression	3
Other data mining techniques	7
TOTAL	74

Table 5. Artificial intelligence techniques used in academic performance.

3.1 Neural Networks

Different studies have carried out the production of academic performance under the use of neural networks. Among the most relevant is a comparative study of several techniques with university students obtained an accuracy of 94.00 %, using demographic and grade variables (Al-Radaideh et al., 2021). On the other hand, the model used a hybrid approach, combining demographic, academic and environmental factors obtaining an accuracy of 85.20 % (Tien et al., 2018).

(Meng et al., 2020) used input (academic performance, personal characteristics, demographic information), output (final exams) variables of 415 students, obtaining an accuracy of 83.10 %.

Similarly (Wang et al. 2022), used a deep learning model to predict the academic performance of college students, with an accuracy of 89.40 %, the variables analyzed are: age, gender, nationality, career, type of admission, high school scores and college entrance exam scores.

(Cui et al., 2022) achieved an accuracy of 86.70 %, variables analyzed: gender, high school grades, college entrance exam scores, and participation in extracurricular activities. (Liang et al., 2022) obtained a high accuracy with a value of 92.40 %, variables under analysis: academic performance, online learning behavior, personal information. (Liu et al., 2022) conducted a study on college students, obtained an accuracy of 90.30 % considering the following variables: academic performance, learning motivation, study habits, time management.

(Wang et al., 2022) combined neural networks with support vector machines, achieving an accuracy of 89.40 % the variables under study were: academic performance, psychological factors, academic motivation, time management. Another study by the same author where he used Long Short-Term Memory (LSTM) neural networks to predict student grades based on variables such as student demographic information, high school grades and class attendance. The model achieved an accuracy of 86.23% in predicting student performance.

3.2 Decision trees

There are several researches that use decision trees, among these we can mention (Guo et al., 2022) who developed a model based on decision trees to predict the academic performance of college students with an accuracy of 89.20 %. The variables analyzed were demographic data, socioeconomic level and academic background. (Zhang et al., 2022) proposed an improved decision tree algorithm to predict the academic performance of



[Licencia Creative Commons Atribución 4.0 Internacional \(CC BY 4.0\)](https://creativecommons.org/licenses/by/4.0/)

college students with a good accuracy of 92.50 %. The variables analyzed were academic performance, daily attendance rate and number of courses completed.

(Wang et al., 2022) developed a decision tree-based method to predict academic performance in undergraduate students with an accuracy of 87.50 %. The variables analyzed were previous academic performance, course attendance rate and socioeconomic status.

(Xiao et al.,2022) proposed an intelligent prediction model for undergraduate academic performance combining decision trees and Bayesian networks with an accuracy of 86.30 %. The variables analyzed were academic background, learning attitude and learning style.

3.3 Neural networks and Decision trees

Several studies combine these techniques, of which we can cite that have higher accuracy we have (Zhang et al., 2019) analyzed the academic performance of Chinese business school students using machine learning algorithms, achieving an accuracy of 90.5 % and highlighting the importance of pre-college factors.

(Khademi et al., 2019) used decision trees and artificial neural network models to predict academic failure, achieving 92.20 % accuracy and identifying attendance and prior GPA (grade point average) as important predictors.

(Alghamdi, 2020) combined these techniques, achieving an accuracy of 86.6 % and highlighting the importance of course-related variables. (Santos, 2020) applied machine learning techniques to predict academic performance, achieving an accuracy of 90.5 % and identifying previous GPA and school type as important predictors.

(Zhang et al., 2020) also combined these techniques to predict students' academic performance, obtaining an accuracy of 86.10 % using academic development variables as well as demographics.

3.4 Support vector machines

(Zhang et al., 2022) aimed to explore the relationship between college students' academic performance and mental health using big data analysis. They collected data from 1800 students in a Chinese university. The variables analyzed included academic performance, mental health, gender, career, and family income. The results showed that there was a significant positive correlation between academic performance and mental health. They also found that main and family income had a significant impact on academic performance, while gender did not. They achieved an accuracy of 88.16%.

4. Discussion

The systematic review is carried out with the purpose of having knowledge of the latest developments according to the line of research, in our case the research that has been developed using artificial intelligence techniques to improve academic performance in higher education institutions. These researches have grown considerably in the last three years, the year where most of these researches have been developed is 2022; this indicates how the interest in this line of research has been increasing.

Of the countries with the greatest number of investigations, we can mention China (n=36), followed by India (n=6) and Spain (n=3), the rest of the countries are just starting in this line of investigation.



[Licencia Creative Commons Atribución 4.0 Internacional \(CC BY 4.0\)](https://creativecommons.org/licenses/by/4.0/)

If we refer to the size of the sample, we can indicate that analyses are carried out with small and large data samples, there is no regulation or standardization in this regard, among these we have that the largest number of investigations are carried out with an interval of 90 - 500 (n=7). Most of the studies do not specify the number of students (n=57).

Regarding the most used artificial intelligence techniques, which give an accuracy higher than 80%, we have artificial neural networks (n=31), followed by decision trees (n=6), it is also important to mention that there are some hybrid methods that also present a high accuracy such as the combination of neural networks and decision trees (n=11).

5. Conclusions and Future Work

Artificial intelligence has undergone significant advances in recent decades, and has been increasingly used to analyze and improve the academic performance of students in higher education institutions, evolving from simple approaches such as machine learning algorithms to more advanced techniques such as neural networks and deep learning algorithms.

Neural network models have been found to be more accurate than decision tree models, and hybrid methods combining several techniques give excellent results. Research has been carried out in a wide range of countries, both those with a long tradition in the use of educational technologies and those that are adopting these technologies more recently. Countries with an outstanding track record in educational research, such as China, have been leaders in the application of artificial intelligence to academic performance in higher education. Other countries, such as India and Spain, have also shown significant interest in this area and have conducted relevant research in the use of artificial intelligence in the academic context.

It has also been found that the most commonly analyzed variables are GPA, previous student performance and participation in extracurricular activities. In addition, the authors have also used different databases and sample populations to conduct their studies.

Being clear about this set of techniques, especially the predictive ones, will allow higher education institutions to develop strategies and policies that allow for student follow-up and monitoring.

In future research, several techniques, such as clustering and autoencoders, can be combined to reduce the dimensionality of the input space and predict students' academic performance. In addition, it would be interesting to explore how these techniques can be adapted to different student populations, such as students of different ages or academic levels. The effectiveness of these techniques in different countries or geographic regions could also be investigated.

Bibliographic references

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.



[Licencia Creative Commons Atribución 4.0 Internacional \(CC BY 4.0\)](https://creativecommons.org/licenses/by/4.0/)

<https://dialnet.unirioja.es/servlet/articulo?codigo=5911340&info=resumen&idoma=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



[Licencia Creative Commons Atribución 4.0 Internacional \(CC BY 4.0\)](https://creativecommons.org/licenses/by/4.0/)

- 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



[Licencia Creative Commons Atribución 4.0 Internacional \(CC BY 4.0\)](https://creativecommons.org/licenses/by/4.0/)

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>

Authors

PATRICIA JIMBO-SANTANA obtained her PhD degree in Computer Science and Security Mathematics from the University of Rovira i Virgili - URV (Spain) with cum laude mention. She obtained her PhD in Computer Science from the National University of La Plata - UNLP (Argentina) in 2020. She obtained a Master's degree in Business Administration from Universidad Andina Simón Bolívar in 2002. Specialist in Finance from Universidad Andina Simón Bolívar in 2000. She obtained a degree in Computer and Information Systems Engineering from the Escuela Politécnica Nacional in 1995.

She is currently a professor at the School of Administrative Sciences of the Universidad Central del Ecuador. She is a computer expert at the Julio Endara Institute of Criminology of the Central University of Ecuador. Her research areas focus on Data Intelligence and its application to data mining. Author of several publications in journals and conferences in the field (Springer, Scopus, Web of Science among others).

LAURA CRISTINA-LANZARINI Scientific Calculator in 1983, Degree in Computer Science in 1986 and PhD in Computer Science in 2017 (University of La Plata - Argentina).

Director of the Master and Specialization in Data Intelligence oriented to Big Data. Since 2014: Secretary of Science and Technique of the Faculty of Computer Science of the UNLP. Since 2001: Full Professor at the Faculty of Computer Science of the UNLP. Since 2000: Visiting Professor at different National Universities of Argentina. His main current research interests are in the field of Data Intelligence and its applications to Signal Processing, Data Mining and Text Mining. Author of numerous publications in journals and conferences in the field.

MONICA JIMBO-SANTANA graduated as a Commercial Engineer in 2004 - Escuela Politécnica del Ejercito (ESPE). Degree in Business Administration from Universidad Central del Ecuador - 2002. Master in Systems Management - Escuela Politécnica del Ejercito (ESPE) -2011, Higher Diploma in University Teaching - Universidad Central del Ecuador -2008. D. in Computer Science from the Universidad de la Plata - Argentina and a PhD in Public Administration from the Universidad Nacional del Litoral - Santa Fe - Argentina.

The lines of research focus on Digital Governance - Electronic Government - Business Intelligence (BI) - Author of numerous publications in scientific journals and conferences. Director of the Business Administration and Public Administration distance learning programs at the Universidad Central del Ecuador - 2015. Professor and researcher at the Faculty of Administrative Sciences of the Central University of Ecuador and Professor of Masters in Computer Systems Audit and Systems Management at the University of the Armed Forces (ESPE) since 2012.



[Licencia Creative Commons Atribución 4.0 Internacional \(CC BY 4.0\)](https://creativecommons.org/licenses/by/4.0/)

MARIO RAÚL-MORALES is a systems engineer graduated from Escuela Politécnica Nacional, Ecuador. His master's degree in business administration was completed at Universidad San Martín de Porres, Peru. He has obtained certifications in project management (PMI) and data analytics, with extensive experience in business projects in the Andean region. D. from the University of Alicante, Spain.

He currently teaches at the Faculty of Engineering and Applied Sciences of the Universidad Central del Ecuador and is a consultant for technology companies in the region. Winner of the "Universidad Central Award for publications of the year 2019". His research interests are related to virtual currencies, blockchain and data analysis.



[Licencia Creative Commons Atribución 4.0 Internacional \(CC BY 4.0\)](https://creativecommons.org/licenses/by/4.0/)