

Article

A Bibliometric Analysis of AI-Driven Performance Prediction in Higher Education

Berat Ujkani ¹, Sead Ujkani ^{2,*}, Daniela Minkovska ³ and Nikolay Hinov ^{4,5,*}

- ¹ Department of Computer Science and Engineering, University “Isa Boletini” Mitrovica, 40000 Mitrovica, Kosovo; berat.ujkani@umib.net
 - ² Department of Accounting, University of Prishtina, 10000 Prishtina, Kosovo
 - ³ Department of Programming and Computer Technologies, Technical University of Sofia, 1000 Sofia, Bulgaria; daniela@tu-sofia.bg
 - ⁴ CoE “National Center of Mechatronics and Clean Technologies”, 1000 Sofia, Bulgaria
 - ⁵ Department of Computer Systems, Faculty of Computer Systems and Technologies, Technical University of Sofia, 1000 Sofia, Bulgaria
- * Correspondence: sead.ujkani@uni-pr.edu (S.U.); hinov@tu-sofia.bg (N.H.)

Abstract

This study presents a comprehensive bibliometric analysis of research publications on artificial intelligence (AI) applications in higher education, with a particular focus on student performance-related studies. Drawing on 1431 documents retrieved from the Web of Science (WoS) Core Collections, advanced bibliometric tools, such as VOSViewer (1.6.19) and Biblioshiny, were used to explore research trends, citation networks, co-authorship patterns, and keyword co-occurrences. The results revealed an increase in AI-related educational research during the COVID-19 pandemic, reflecting the rise in the reliance on AI for enhancing learning outcomes in times of educational disruption. The findings also indicated that the research focus is gradually moving towards using AI for educational assessment, emphasizing the importance of accurate and data-driven evaluation of student performance. The co-occurrence of keywords and citation analyses confirmed that machine learning, deep learning, and predictive modeling are among the dominant AI techniques applied to assess and predict student outcomes. Furthermore, the study highlights AI’s potential in identifying learning gaps and enabling personalized interventions, allowing educators to address students’ specific needs more effectively. This new trend suggests a growing recognition of AI’s role in refining educational methodologies and improving performance evaluations at the tertiary level.

Keywords: artificial intelligence; student performance prediction; bibliometric analysis; educational assessment; higher education



Academic Editors: Petros Lameris, Sylvester Arnab and Panagiotis Petridis

Received: 27 June 2025

Revised: 14 August 2025

Accepted: 16 August 2025

Published: 22 August 2025

Citation: Ujkani, B.; Ujkani, S.; Minkovska, D.; Hinov, N. A Bibliometric Analysis of AI-Driven Performance Prediction in Higher Education. *Information* **2025**, *16*, 713. <https://doi.org/10.3390/info16090713>

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Growth in technology has changed how data/information is collected and analyzed. Nowadays, data is systematically aggregated, analyzed, and interpreted to inform practices, methodologies, and decision-making across diverse sectors [1], including the education sector [2,3]. Thus, there is a growing research interest in the analysis of data generated through educational technologies recently [4], owing to its potential to drive innovation in pedagogical practices [5] and to challenge established educational paradigms [6]. Concurrently, the proliferation of sophisticated analytical tools designed for academic data analysis has significantly increased [7]. This evolution has prompted educators to critically

reassess traditional conceptions of teaching and learning, encouraging a shift towards data-informed instructional strategies that enhance educational outcomes.

The International Educational Data Mining (EDM) Society emphasized that EDM involves developing methodologies for analyzing the diverse and intricate data generated within educational environments to enhance our understanding of student behaviours and learning contexts [8]. Before now, researchers have recognized the potential for longitudinal tracking of student learning progress, which enables the identification of activities that significantly contribute to improved learning outcomes [9]. Integrating big data technologies into educational settings has led to an important shift from traditional, instructor-centered pedagogies to more learner-centric models, largely driven by advancements in e-learning. EDM is specifically defined as the application of data mining techniques to educational data, focusing on addressing critical questions and challenges within the educational landscape [10]. The EDM framework informs instructional design as well as empowers educators to make data-driven decisions that enhance student engagement and achievement.

In practice, EDM entails analyzing large data volumes generated by the integration of technology in education, such as data mining, statistical modeling, and artificial intelligence (AI). It aims to address educational queries through computational methods, analyzing and interpreting educational content [11]. The EDM field enables a profound transformation of the education system, urging a reevaluation of its structure to elevate a nation's standing globally [12]. Using data mining, the primary technique behind EDM, institutions and educators can uncover hidden patterns and insights from various information sources. This will help in facilitating a comprehensive review of student performance and related factors. Employing different algorithms allows for the detailed categorization of student databases based on multiple factors and to prioritize psychological over intellectual traits, offering insights into student behavior patterns, helping to effectively profile students in any educational environment [13,14].

Bibliometric analysis serves as a comprehensive methodology for both quantitative and qualitative assessment of academic publications, and it has been extensively utilized across various scientific domains [15]. This approach enables researchers to conduct a comprehensive examination of topic evolution, identify core research methodologies, and pinpoint emerging areas for further inquiry within a specific field. It is particularly well suited for unravelling the knowledge structure and developmental direction of studies focused on student behaviours about their academic objectives [16,17].

Bibliometric analysis uses statistical and visual techniques to illustrate the knowledge structures and developmental trends associated with a given topic. By analyzing publication data, it is easier to identify prevailing research trends and emergent themes. Additionally, it facilitates the evaluation of publication outputs by authors and institutions, allowing for the mapping of international collaboration networks and geographical distribution patterns [18]. Also, this comprehensive analytical framework enhances the understanding of scholarly communication as well as informs strategic decision-making in educational research and policy development.

There are diverse bibliometric analysis tools, each possessing unique advantages and limitations. It is, therefore, prudent to employ multiple tools for a comprehensive analysis. Among the most widely utilized software packages for bibliometric and visualization analysis are CiteSpace, Visualization of Similarities Viewer (VOSviewer), and R-bibliometrix [19,20]. These tools facilitate the collection of important information from included documents, enable the identification of relevant articles within cooperative networks, and allow for the assessment of contributions from authors, institutions, and countries or regions [21].

Raw bibliometric data can be obtained from different databases, including WoS, PubMed, and Scopus. WoS and Scopus are extensive subscription-based databases, while

PubMed serves as a free resource primarily focused on the biomedical literature, positioning it as the premier database for electronic biomedical research. Scopus encompasses a broader range of journals, which is beneficial for conducting keyword searches and citation analyses. However, it is worth noting that Scopus is limited to publications from 1992 onwards, whereas WoS provides access to materials published before 1992 [22]. This temporal breadth in WoS may offer a more comprehensive perspective for historical bibliometric studies, further emphasizing the importance of selecting appropriate databases based on research needs.

This paper presents a bibliometric study aimed at revealing state-of-the-art research works on AI applications in predicting student performance within higher education. To the best of our knowledge, this represents a novel endeavor in providing a comprehensive bibliometric review of AI-driven performance prediction in the educational sector. Furthermore, the paper highlights the significant influence of the COVID-19 pandemic on the accelerated adoption of AI technologies in educational contexts.

The major contributions of the paper are as follows:

1. This study provides an extensive bibliometric overview of the literature on AI applications related to predicting student performance in higher education, mapping publication trends, influential authors, and research collaborations.
2. The paper identifies key trends in AI research within higher education, including emerging themes, methodologies, and technologies that dominate the field.
3. It elucidates the critical role of the COVID-19 pandemic in driving the adoption of AI in educational settings, reflecting on how this global event has reshaped educational practices.
4. The study evaluates the publication output from various authors, institutions, and countries, providing insights into collaborative networks and the geographical distribution of research efforts.
5. By identifying gaps in the existing literature, the paper proposes avenues for future research, encouraging further exploration of hybrid methodologies that combine AI with traditional educational assessment techniques.

2. Methodology

The first step of a bibliometric study is to collect data, which is then presented in bibliometric and statistical ways. Various methods exist for obtaining raw data. For this study, plain text files of the raw data were obtained from the WoS platform, while VOSviewer and the R tool were utilized to map the data.

2.1. Search Strategy and Data Collection

As stated above, the WoS platform was utilized to collect raw data, specifically the WoS Core Collection. The keywords used for the search are given in Table 1. The asterisk (*) indicates truncation, allowing the query to capture plural forms and different suffixes of the same root term.

Table 1. WoS-compatible query string for bibliometric analysis.

Query String
“academic performance*” OR “education modeling” OR “pedagogy*” OR “educational data*” OR “learning analytics*” OR “academic success*” OR “student success*” OR “student dropout*” OR “student enrollment*” OR “precision education*” OR “predictive learning*” OR “student academic*” AND “higher education*” OR “university” OR “universities” OR “tertiary education”

A total of 1431 records were retrieved from the following databases: Science Citation Index Expanded (SCIE), Social Sciences Citation Index (SSCI), Conference Proceedings Citation Index—Science (CPCI-S), and Emerging Sources Citation Index (ESCI). The publication time frame was established from January 2009 to December 2023, encompassing a targeted span of 15 years.

For subsequent analyses, the raw data files were exported in “Plain text file” format. Due to the limitation of exporting only 500 records when selecting “Full records and cited references” from the “Record content” options, the export process was conducted in three batches: files 1–500, 501–1000, and 1001–1431. Following this, a Python script (CPython 3.10.13) was employed to merge the exported plain text files into a single comprehensive file containing all records for further analysis.

2.2. Data Extraction and Analysis

After integrating all the files into a single file, the data were analyzed, and the relevant information was extracted using the R tool (R 4.1.2), Biblioshiny (R 4.1.2), and VOSviewer (1.6.19). Specifically, VOSviewer was used to generate and evaluate bibliometric networks such as journals, researchers, or individual articles based on citation, bibliographic coupling, co-citation, or co-authorship links. VOSviewer was also used for text mining capabilities, such as constructing and visualizing co-occurrence networks of phrases extracted from the scientific literature. R-bibliometrix (R 4.1.2) was used for quantitative analysis.

This study uses data exclusively from the Web of Science database, which may limit the inclusion of research from underrepresented sources.

3. Results and Discussion

This section presents the results of the bibliometric experiments conducted, including the use of word cloud, author, citation, and co-author mapping. In addition, the section presents the network visualizations of citation documents, citation sources, and network visualizations revealing meaningful insights into the trends and research direction on AI-driven approaches to education.

3.1. Basic Information on Publications

Using the search criteria shown above, a total of 1431 documents were used for further analysis. The range of the era was from 2009 to 2023. There were 802 sources involved in the publication of these items, representing an annual growth of 31.97%. There was a total of 4250 authors who have contributed to these publications, with an average citation frequency of 9.57% per document. In Table 2, a comprehensive description is presented.

Table 2. Main information of the extracted documents.

Description	Results
Timespan	2009–2023
Sources (journals, books, etc.)	802
Documents	1431
Annual growth rate %	31.97
Document average age	4.15
Average citations per doc	9.573
References	40.188
Document contents	
Keywords plus (ID)	902
Author’s keywords (DE)	3234

Table 2. Cont.

Description	Results
Authors	
Authors	4250
Authors of single-authored docs	132
Author collaboration	
Single-authored docs	137
Co-authors per doc	3.55
International co-authorships %	23.83
Document types	
Article	637
Article; book chapter	1
Article; data paper	5
Article; early access	206
Article; proceedings paper	2
Correction	1
Editorial material	5
Proceedings paper	515
Proceedings paper; early access	2
Review	39
Review; early access	18

3.2. Annual Production of Articles

The data is based on the documents published from 2009 to 2023. Only five articles were published in 2009. The creation of articles was growing over time. There was a significant growth in literature production from 2018 to 2022. In 2022, it reached its peak, and a total of 257 documents were published, but there was a decrease in the production of papers in 2023, with only 243 documents published. Trends in production over time are provided in Figure 1.

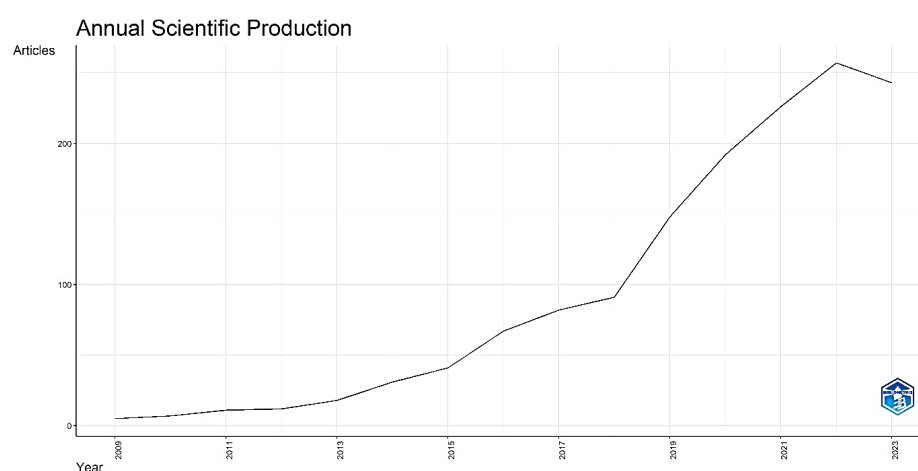


Figure 1. Annual production of documents (2009–2023).

3.3. Most Relevant Sources

It was found that a total of 802 sources were involved in the publication of the entire literature. The data was analyzed to find out the top ten most relevant sources based on several publications, as shown in Figure 2.

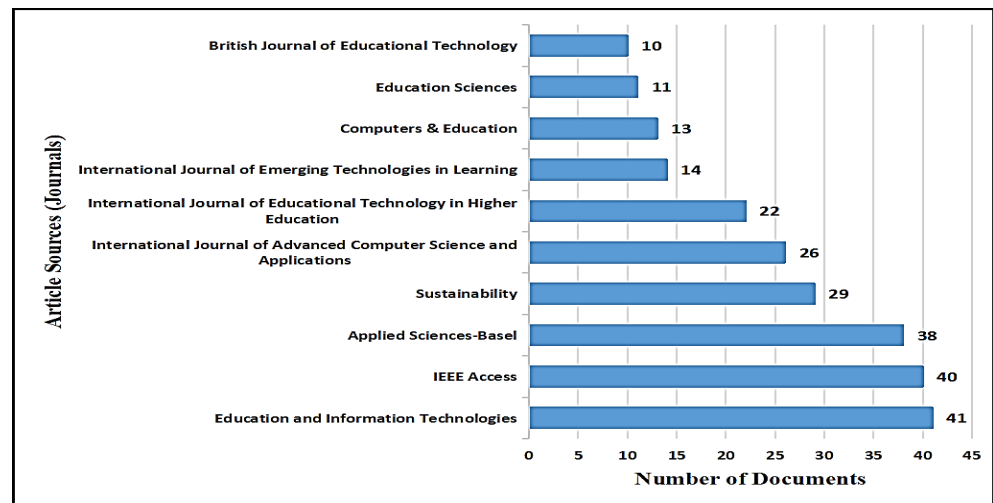


Figure 2. Top 10 relevant sources.

3.4. Most Relevant Authors and Institutions

A total of 4250 authors have contributed to publishing this literature. The data was analyzed to find the most relevant authors according to the number of publications. Kotsiantis S stood on the top, with a total of 13 publications. Figure 3 shows the top ten most relevant authors.

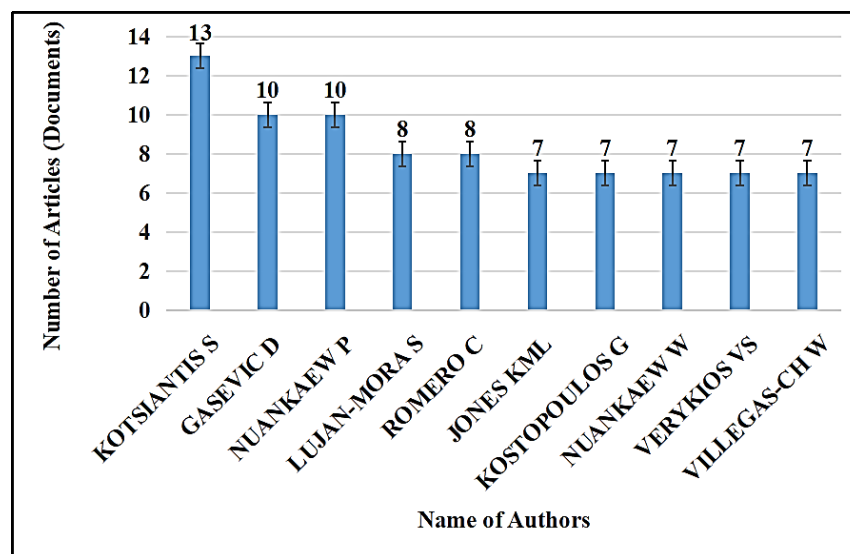


Figure 3. Top 10 relevant authors.

Several institutions were involved in the research. The top five institutions based on the number of publications are shown in Table 3. Tecnologico De Monterrey appeared on top, having published 40 articles.

Table 3. The top five institutions with the largest number of publications.

Affiliation	Articles
Tecnologico De Monterrey	40
King Abdulaziz University	32
Egyptian Knowledge Bank	18
California State University	14
University of Patras	13

Similarly, the data was analyzed to find the corresponding author's countries. In Table 4, the top ten countries based on the number of publications are shown. These results imply that Africa is yet to be fully represented in AI for education research. This is understandable owing to the need for an increase in education funding, as the countries revealed here (China and the USA) are known to have better education funding strategies.

Table 4. Corresponding author's countries.

Country	Articles	SCP	MCP	Freq	MCP_Ratio
China	182	147	35	0.127	0.192
USA	145	123	22	0.101	0.152
Spain	81	66	15	0.057	0.185
India	79	73	6	0.055	0.076
Saudi Arabia	59	42	17	0.041	0.288
Australia	56	35	21	0.039	0.375
United Kingdom	54	39	15	0.038	0.278
Germany	37	24	13	0.026	0.351
Brazil	36	29	7	0.025	0.194
Mexico	35	28	7	0.024	0.2

3.5. Most Cited Documents

Citing a document by the researchers is usually based on its usefulness to the scientific community as well as to the public. A most cited document means it has more impact [23]. The top 10 most cited documents from the extracted and analyzed data are shown in Table 5.

Table 5. Top ten most cited articles.

Paper	DOI	Total Citations	TC Per Year	Normalized TC
Zawacki-Richter O., 2019 [24]	https://doi.org/10.1186/s41239-019-0171-0	403	67.17	22.82
Asif R., 2017 [10]	https://doi.org/10.1016/j.compedu.2017.05.007	223	27.88	15.59
Aldowah H., 2019 [25]	https://doi.org/10.1016/j.tele.2019.01.007	192	32.00	10.87
Costa Eb., 2017 [26]	https://doi.org/10.1016/j.chb.2017.01.047	184	23.00	12.86
Cerezo R., 2016 [27]	https://doi.org/10.1016/j.compedu.2016.02.006	172	19.11	12.31
Onan A., 2021 [28]	https://doi.org/10.1002/cae.22253	167	41.75	21.51
Romero C., 2013 [29]	https://doi.org/10.1002/cae.20456	162	13.50	6.37
Waheed H., 2020 [30]	https://doi.org/10.1016/j.chb.2019.106189	157	31.40	11.62
Alyahyan E., 2020 [31]	https://doi.org/10.1186/s41239-020-0177-7	128	25.60	9.47
Natek S., 2014 [32]	https://doi.org/10.1016/j.eswa.2014.04.024	110	10.00	7.65

3.6. Core Sources by Bradford's Law

The core sources by Bradford's Law (Bradford's Law is foundational in the field of bibliometrics and is used to identify core journals most relevant to a specific field. It helps reveal that a few journals contain the majority of significant articles on a subject, while the

remainder are scattered across many publications) have been identified. Table 6 provides a brief description of the core sources.

Table 6. Core sources by Bradford’s Law.

Source	Rank	Freq	cumFreq	Zone
<i>Education and Information Technologies</i>	1	41	41	Zone 1
<i>IEEE Access</i>	2	40	81	Zone 1
<i>Applied Sciences-Basel</i>	3	38	119	Zone 1
<i>Sustainability</i>	4	29	148	Zone 1
<i>International Journal of Advanced Computer Science and Applications</i>	5	26	174	Zone 1
<i>International Journal of Educational Technology in Higher Education</i>	6	22	196	Zone 1
<i>International Journal of Emerging Technologies in Learning</i>	7	14	210	Zone 1
<i>Computers & Education</i>	8	13	223	Zone 1
<i>Education Sciences</i>	9	11	234	Zone 1
<i>British Journal of Education Technology</i>	10	10	244	Zone 1

It is noteworthy that most of the core sources are either open-access or hybrid journals. This confirms the established argument for open access articles, and it gives the following advantages: (i) over 50% chance of more citations, (ii) seven times higher chance of being downloaded, and (iii) faster means to make research works available to the least developed countries where funding is limited.

3.7. WordCloud Map and Trending Topics

The data was analyzed to build a WordCloud map. The most frequently used word was “performance”, with a frequency of 132. The second most frequently used word was “higher-education”, with a frequency of 100. Figure 4 shows the map.



Figure 4. Word cloud based on author keywords.

In addition, the trending topics over time have been mapped, as shown in Figure 5. Terms like *performance*, *students*, and *higher education* have gained prominence in recent

years, while the earlier focus was on *systems* and *scores*. The size of the bubbles indicates how frequently each term appeared.

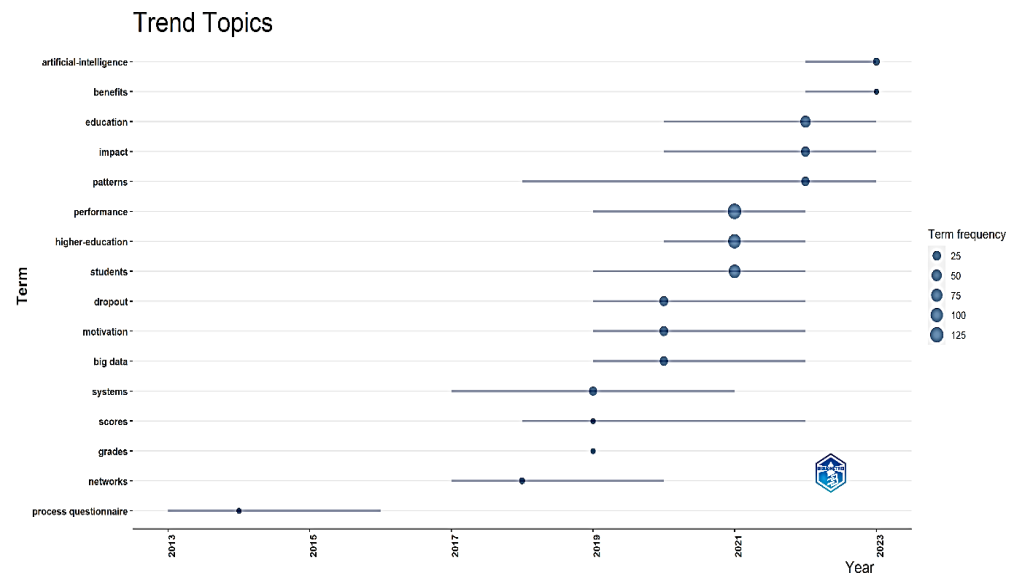


Figure 5. Trending topics over time.

3.8. Co-Authorship—Authors

The co-authorship analysis employed the “Full Counting” method. The selection criteria for authors included a minimum of five published documents and a minimum of 15 citations per author. Out of 4347 authors, 23 satisfied these thresholds. The resulting network visualization, as depicted in Figure 6, demonstrates a dispersed connection pattern.

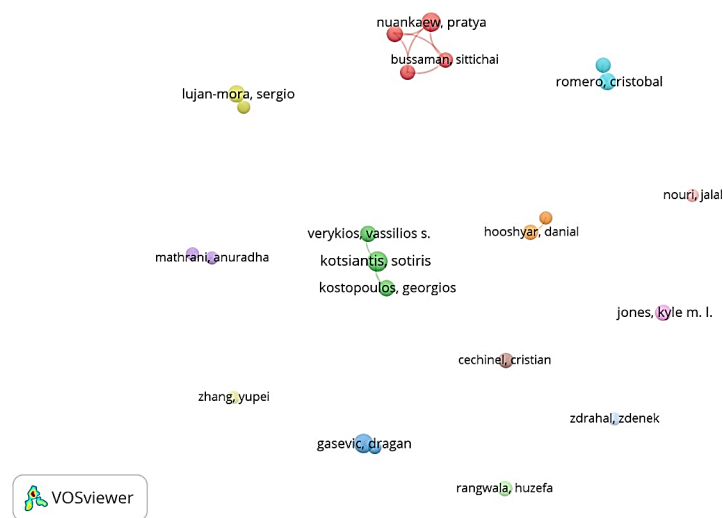


Figure 6. Co-authorship—author map.

Subsequently, the co-occurrence network of all keywords was generated using the same method, preserving the keywords in their original form. The minimum occurrence threshold for each keyword was set to 15. Out of 3918 keywords, 86 met this criterion.

3.9. Citations

Citations were analyzed and mapped across five categories: (A) Documents, (B) Sources, (C) Authors, (D) Organizations, and (E) Countries. A minimum citation

threshold of 15 per document was again applied. Out of 1435 documents, 256 met this criterion. The corresponding network visualization is depicted in Figure 7.

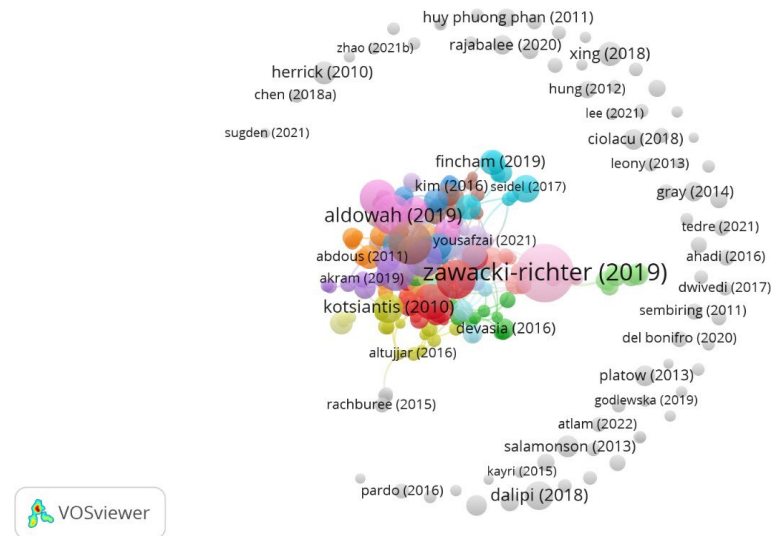


Figure 7. Citations—documents network visualization.

Similarly, citations to sources were mapped using specific inclusion criteria. Each source included in the analysis was required to have published at least five documents, with a minimum of five citations per source. Out of 800 total sources, 43 met these thresholds. The resulting network visualization is presented in Figure 8.

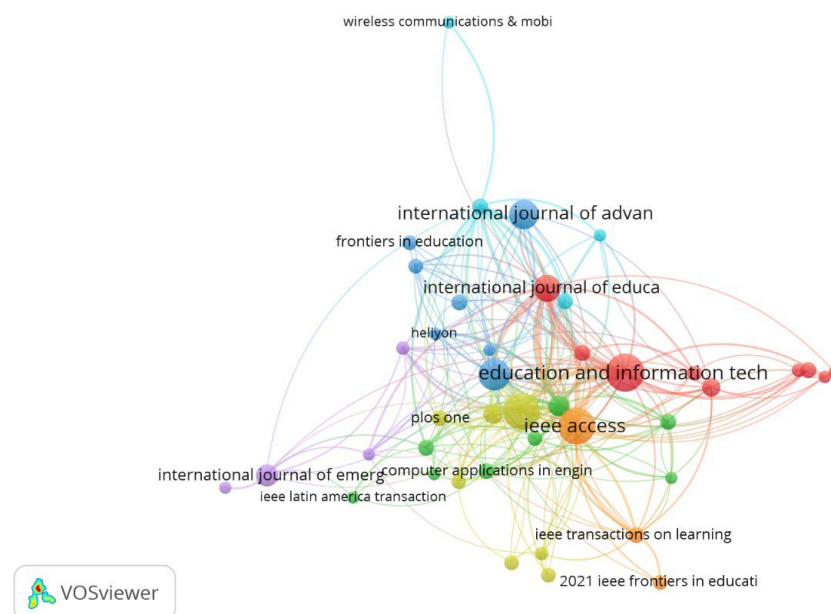


Figure 8. Citations—sources mapping.

Furthermore, the citation–author nodes were analyzed. Since the focus was not on the number of authors in a document, the number of authors was ignored. In addition, the conversion of the author names to initials was also ignored. Each author should have at least five documents and 15 citations. Of 4347 authors, 23 met this criterion. Network visualization was created, as shown in Figure 9. The visualization reveals a fragmented citation landscape, where top-cited authors, such as Kotsiantis Sotiris, appear as central but largely unconnected nodes, indicating minimal co-citation among prolific contributors.

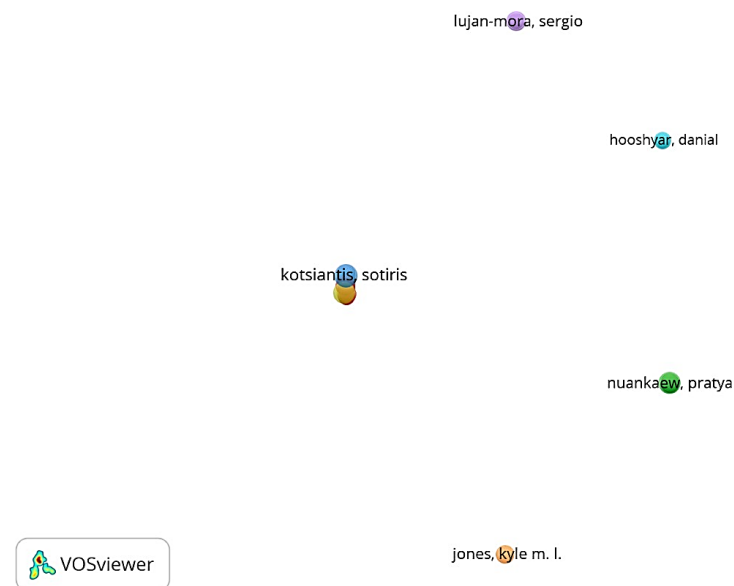


Figure 9. Network visualization of citation–author nodes.

In addition, a citation network for organizations was mapped. Documents co-authored by a large number of organizations were excluded from the analysis. The maximum number of documents per organization was limited to five. To form the network, each organization was required to have a minimum of five documents and 10 citations. Out of 1723 organizations, 74 met these criteria. Network visualization was generated for the largest connected component, consisting of 67 organizations, as depicted in Figure 10.

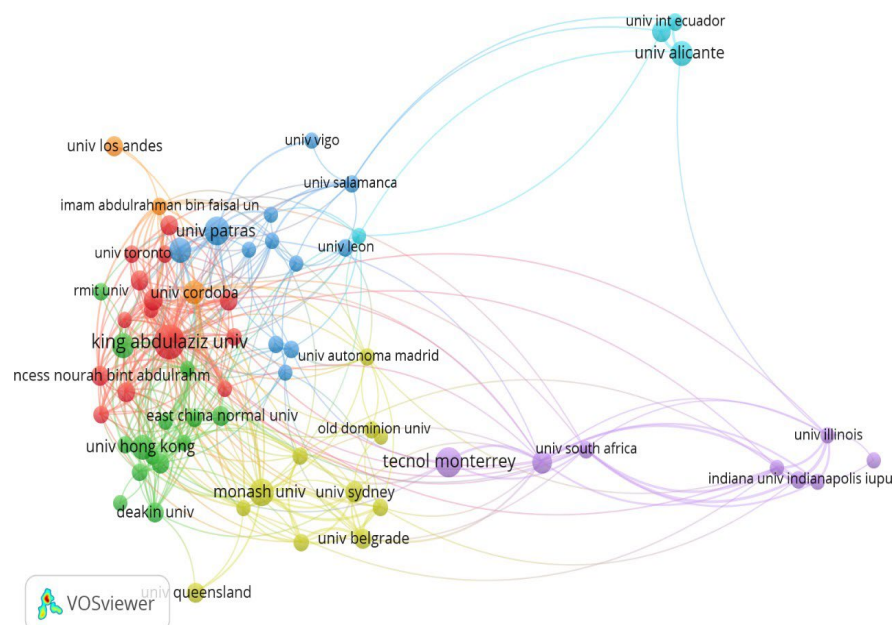


Figure 10. Network visualization of citation–organization nodes.

Lastly, citation–country nodes were mapped. The large documents co-authored by a large number of countries were ignored. The minimum number of documents of a country was chosen to be 10, and the minimum number of citations was chosen to be 20.

Of 101 countries, 47 met these thresholds. These 47 countries were mapped, and a network visualization is shown in Figure 11.

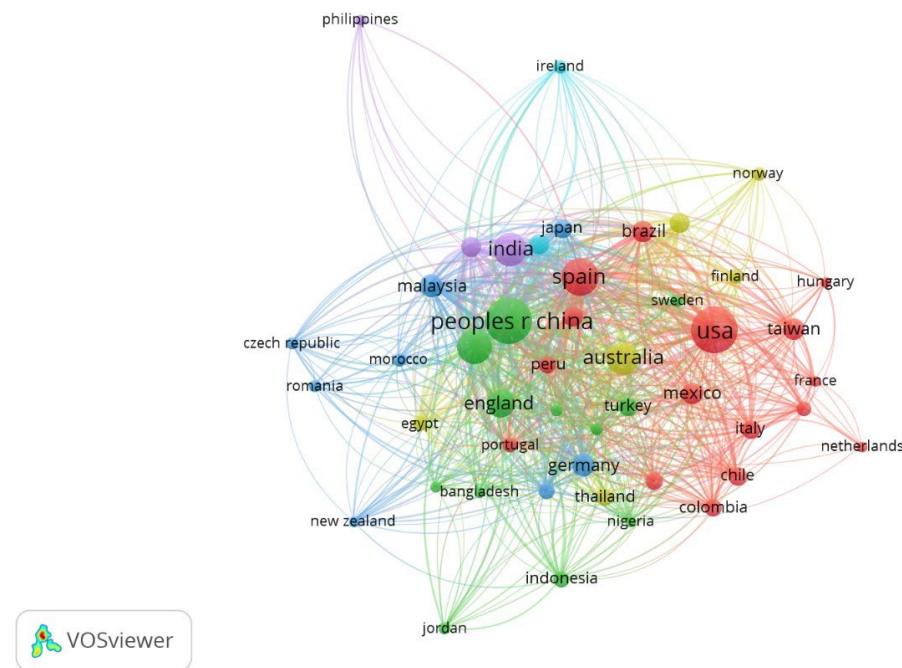


Figure 11. Network visualization of citation–country nodes.

3.10. Bibliographic Coupling

Bibliographic coupling was performed by targeting (A) Documents, (B) Sources, (C) Authors, (D) Organizations, and (E) Countries [24]. Corresponding nodes were mapped with full counting. The minimum number of citations of a document was set to be 10. Of 1435 documents, 362 met this criterion. A network visualization map was generated of the 350 connected items, as shown in Figure 12.

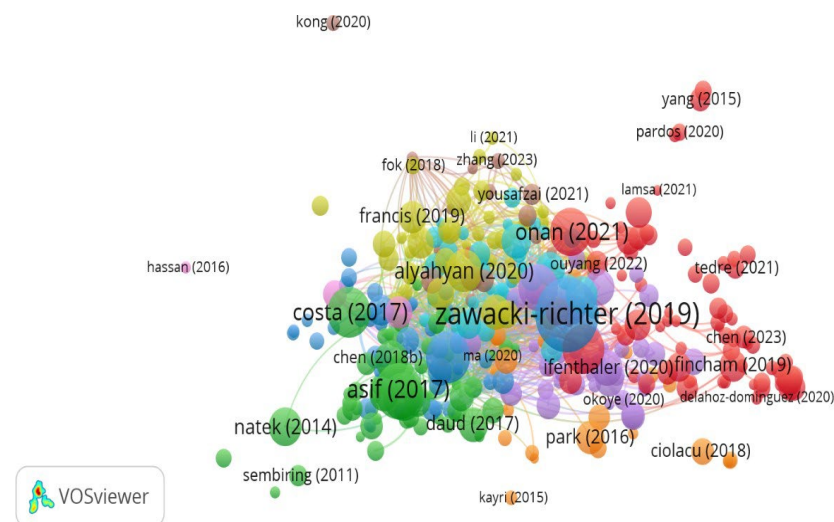


Figure 12. Network visualization of bibliographic coupling–document nodes.

Similarly, bibliographic sources were mapped to generate a network visualization. The criteria established included a minimum of five published documents and at least five citations per source. Out of 800 sources analyzed, 43 met these thresholds. Bibliographic authors were also mapped to create a network visualization. The criteria established included a minimum of five published documents and at least 20 citations per author. Out of a total of 4547 authors, 22 met these thresholds.

The findings reveal several thematic clusters related to AI techniques and educational outcomes, as well as active collaborations among researchers and institutions. While the analysis does not assess the pedagogical effectiveness of AI directly, it does underscore the field’s evolution toward more diverse and complex applications of AI in education. The research highlights the increasing relevance of AI in the educational sector and also suggests that future work should explore hybrid AI methodologies, such as combining machine learning with advanced statistical models, to further improve predictive accuracy and the personalization of educational experiences.

As a recommendation to stakeholders championing an AI-driven approach to education, it is recommended that dedicated funding, especially in least developed nations, will be critical to the full realization of this goal.

Author Contributions: B.U., S.U., D.M., and N.H. were involved in the full process of producing this paper, including conceptualization, methodology, modeling, validation, visualization, and preparing the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the European Regional Development Fund under the “Research Innovation and Digitization for Smart Transformation” program 2021–2027 under Project BG16RFPR002-1.014-0006 “National Centre of Excellence Mechatronics and Clean Technologies”, and the APC was funded by Project BG16RFPR002-1.014-0006.

Institutional Review Board Statement: Not applicable.

Data Availability Statement: The data for this study were obtained from the Web of Science database (<https://www.webofscience.com> (accessed on 15 September 2024) and are subject to access restrictions.

Acknowledgments: The authors thank the anonymous reviewers for their insightful comments and constructive suggestions, which substantially improved the manuscript.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

Notation	Meaning
AI	Artificial Intelligence
CPCI-S	Conference Proceedings Citation Index—Science
DOI	Digital Object Identifier
EDM	Educational Data Mining
ESCI	Emerging Sources Citation Index
MCP	Multiple-Country Publication
MCP_Ratio	Multiple-Country Publication Ratio
R-Tool	R-Package
SCIE	Science Citation Index Expanded
SCP	Single-Country Publication
SSCI	Social Science Citation Index
TC	Total Citations
VoSviewer	Visualization of Similarities Viewer
WoS	Web of Science

References

1. Kar, A.K.; Dwivedi, Y.K. Theory building with big data-driven research—Moving away from the “What” towards the “Why”. *Int. J. Inf. Manag.* **2020**, *54*, 102205. [CrossRef]

2. Baek, C.; Doleck, T. A Bibliometric Analysis of the Papers Published in the Journal of Artificial Intelligence in Education from 2015–2019. *Int. J. Learn. Anal. Artif. Intell. Educ. (ijAI)* **2020**, *2*, 67–84. [CrossRef]

3. Charitopoulos, A.; Rangoussi, M.; Koulouriotis, D. On the Use of Soft Computing Methods in Educational Data Mining and Learning Analytics Research: A Review of Years 2010–2018. *Int. J. Artif. Intell. Educ.* **2020**, *30*, 371–430. [\[CrossRef\]](#)
4. Ang, K.L.M.; Ge, F.L.; Seng, K.P. Big Educational Data & Analytics: Survey, Architecture and Challenges. *IEEE Access* **2020**, *8*, 116392–116414. [\[CrossRef\]](#)
5. Doleck, T.; Lemay, D.; Basnet, R.; Bazelais, P. Predictive analytics in education: A comparison of deep learning frameworks. *Educ. Inf. Technol.* **2020**, *25*, 1951–1963. [\[CrossRef\]](#)
6. Bakti, I.K.; Zulkarnain; Yaran, A.; Rusdi; Syaifudin, M.; Syafaq, H. The Role of Artificial Intelligence in Education: A Systematic Literature Review. *J. Iqra Kaji. Ilmu Pendidik.* **2023**, *8*, 182–197. [\[CrossRef\]](#)
7. Rimpay; Dhankhar, A.; Solanki, K. Educational Data Mining tools and Techniques used for Prediction of Student's Performance: A Study. In Proceedings of the 2022 10th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO), Noida, India, 13–14 October 2022; pp. 1–5. [\[CrossRef\]](#)
8. Du, X.; Yang, J.; Hung, J.L.; Shelton, B. Educational data mining: A systematic review of research and emerging trends. *Inf. Discov. Deliv.* **2020**, *48*, 225–236. [\[CrossRef\]](#)
9. Roediger, H.L.; Karpicke, J.D. Test-Enhanced Learning: Taking Memory Tests Improves Long-Term Retention. *Psychol. Sci.* **2006**, *17*, 249–255. [\[CrossRef\]](#)
10. Asif, R.; Merceron, A.; Ali, S.A.; Haider, N.G. Analyzing undergraduate students' performance using educational data mining. *Comput. Educ.* **2006**, *113*, 177–194. [\[CrossRef\]](#)
11. Boztas, G.D.; Berigel, M.; Altınay, F. A bibliometric analysis of Educational Data Mining studies in global perspective. *Educ. Inf. Technol.* **2023**, *29*, 8961–8985. [\[CrossRef\]](#)
12. Etherington, M. The Challenge with Educational Transformation. *J. Cult. Values Educ.* **2019**, *2*, 96–112. [\[CrossRef\]](#)
13. Angeline, D.M.D. Association Rule Generation for Student Performance Analysis using Apriori Algorithm. *SIJ Trans. Comput. Sci. Eng. Its Appl. CSEA* **2013**, *1*, 16–20. [\[CrossRef\]](#)
14. Sinaga, K.P.; Yang, M.S. Unsupervised K-Means Clustering Algorithm. *IEEE Access* **2020**, *8*, 80716–80727. [\[CrossRef\]](#)
15. Mudawi, N.A.; Pervaiz, M.; Alabdullah, B.I.; Alazeb, A.; Alshahrani, A.; Alotaibi, S.S.; Jalal, A. Predictive Analytics for Sustainable E-Learning: Tracking Student Behaviors. *Sustainability* **2023**, *15*, 14780. [\[CrossRef\]](#)
16. Prahani, B.K.; Rizki, I.A.; Jatmiko, B.; Suprpto, N.; Tan, A. Artificial Intelligence in Education Research During The Last Ten Years: A Review and Bibliometric Study. *Int. J. Emerg. Technol. Learn. IJET* **2022**, *17*, 169–188. [\[CrossRef\]](#)
17. Akhmadieva, R.S.; Udina, N.N.; Kosheleva, Y.P.; Zhdanov, S.P.; Timofeeva, M.O.; Budkevich, R.L. Artificial intelligence in science education: A bibliometric review. *Contemp. Educ. Technol.* **2023**, *15*, ep460. [\[CrossRef\]](#)
18. Choudhri, A.F.; Siddiqui, A.; Khan, N.R.; Cohen, H.L. Understanding Bibliometric Parameters and Analysis. *RadioGraphics* **2015**, *35*, 736–746. [\[CrossRef\]](#)
19. Eck, N.J.V.; Waltman, L. Software survey: VOSviewer, a computer program for bibliometric mapping. *Scientometrics* **2010**, *84*, 523–538. [\[CrossRef\]](#)
20. Dai, Z.; Xu, S.; Wu, X.; Hu, R.; Li, H.; He, H.; Hu, J.; Liao, X. Knowledge Mapping of Multicriteria Decision Analysis in Healthcare: A Bibliometric Analysis. *Front. Public Health* **2022**, *10*, 895552. [\[CrossRef\]](#)
21. Devos, P.; Menard, J. Bibliometric analysis of research relating to hypertension reported over the period 1997–2016. *J. Hypertens.* **2019**, *37*, 2116–2122. [\[CrossRef\]](#)
22. Falagas, M.E.; Pitsouni, E.I.; Malietzis, G.A.; Pappas, G. Comparison of PubMed, Scopus, Web of Science, and Google Scholar: Strengths and weaknesses. *FASEB J.* **2008**, *22*, 338–342. [\[CrossRef\]](#)
23. Tikhonova, E.; Raitskaya, L. Citations and References: Guidelines on Literature Practices. *J. Lang. Educ.* **2022**, *8*, 5–10. [\[CrossRef\]](#)
24. Zawacki-Richter, O.; Marín, V.I.; Bond, M.; Gouverneur, F. Systematic review of research on artificial intelligence applications in higher education—where are the educators? *Int. J. Educ. Technol. High. Educ.* **2019**, *16*, 39. [\[CrossRef\]](#)
25. Aldowah, H.; Al-Samarraie, H.; Fauzy, W.M. Educational data mining and learning analytics for 21st century higher education: A review and synthesis. *Telemat. Inform.* **2019**, *37*, 13–49. [\[CrossRef\]](#)
26. Costa, E.B.; Fonseca, B.; Santana, M.A.; de Araújo, F.F.; Rego, J. Evaluating the effectiveness of educational data mining techniques for early prediction of students' academic failure in introductory programming courses. *Comput. Hum. Behav.* **2017**, *73*, 247–256. [\[CrossRef\]](#)
27. Cerezo, R.; Sánchez-Santillán, M.; Paule-Ruiz, M.P.; Núñez, J.C. Students' LMS interaction patterns and their relationship with achievement: A case study in higher education. *Comput. Educ.* **2016**, *96*, 42–54. [\[CrossRef\]](#)
28. Onan, A. Sentiment analysis on massive open online course evaluations: A text mining and deep learning approach. *Comput. Appl. Eng. Educ.* **2021**, *29*, 572–589. [\[CrossRef\]](#)
29. Romero, C.; Espejo, P.G.; Zafra, A.; Romero, J.R.; Ventura, S. Web usage mining for predicting final marks of students that use Moodle courses. *Comput. Appl. Eng. Educ.* **2013**, *21*, 135–146. [\[CrossRef\]](#)
30. Waheed, H.; Hassan, S.-U.; Aljohani, N.R.; Hardman, J.; Alelyani, S.; Nawaz, R. Predicting academic performance of students from VLE big data using deep learning models. *Comput. Hum. Behav.* **2020**, *104*, 106189. [\[CrossRef\]](#)

31. Alyahyan, E.; Düşteğör, D. Predicting academic success in higher education: Literature review and best practices. *Int. J. Educ. Technol. High. Educ.* **2020**, *17*, 3. [[CrossRef](#)]
32. Natek, S.; Zwilling, M. Student data mining solution—knowledge management system related to higher education institutions. *Expert Syst. Appl.* **2014**, *41*, 6400–6407. [[CrossRef](#)]
33. Guo, H.; Shen, Z.; Zeng, J.; Hong, N. Hybrid Methods of Bibliographic Coupling and Text Similarity Measurement for Biomedical Paper Recommendation. In Proceedings of the MEDINFO 2021: One World, One Health—Global Partnership for Digital Innovation, Virtual Event, 2–4 October 2022; IOS Press: Amsterdam, The Netherlands, 2022; pp. 287–291. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.