frontiers | Frontiers in Education

TYPE Systematic Review
PUBLISHED 25 September 2025
DOI 10.3389/feduc.2025.1651484

Check for updates

OPEN ACCESS

EDITED BY
Nada Mallah Boustani,
Saint Joseph University, Lebanon

REVIEWED BY
Nadezhda Kalmazova,
Peoples' Friendship University of Russia, Russia
Richard Sunarjo,
Raharja University, Indonesia
Surabhi Singh,
O.P. Jindal Global University, India

*CORRESPONDENCE
Gökmen Dağlı
✉ gokmen.dagli@kyrenia.edu.tr

RECEIVED 01 July 2025
ACCEPTED 05 September 2025
PUBLISHED 25 September 2025

CITATION
Alpürk K, Ayaz A, Altınay F, Altınay Z, Berigel DS and Dağlı G (2025) Artificial intelligence applications in entrepreneurship and online education: insights from bibliometric and topic modeling analyses. *Front. Educ.* 10:1651484.
doi: 10.3389/feduc.2025.1651484

COPYRIGHT
© 2025 Alpürk, Ayaz, Altınay, Altınay, Berigel and Dağlı. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Artificial intelligence applications in entrepreneurship and online education: insights from bibliometric and topic modeling analyses

Kemal Alpürk[1], Ahmet Ayaz[2], Fahriye Altınay[3], Zehra Altınay[4], Duygu Solak Berigel[5] and Gökmen Dağlı[6]*

[1]Faculty of Education, Near East University, Nicosia, Cyprus, [2]Digital Transformation Office, Karadeniz Technical University, Trabzon, Türkiye, [3]Faculty of Education, Societal Research and Development Center, Institute of Graduate Studies, Near East University, Nicosia, Cyprus, [4]Faculty of Education, Societal Research and Development Center, Near East University, Nicosia, Cyprus, [5]Department of Computer Technology, Çarşıbaşı Vocational School, Trabzon University, Trabzon, Türkiye, [6]Faculty of Education, University of Kyrenia, Kyrenia, Cyprus

Introduction: The convergence of AI, entrepreneurship, and online education has accelerated, yet their intersection remains under-mapped.
Methods: We analyzed 489 peer-reviewed articles (2010–2024) from Scopus and Web of Science using Bibliometrix for bibliometrics and LDA topic modeling (Gensim/pyLDAvis).
Results: Five thematic clusters emerged; "Teaching and Pedagogical Innovations" and "Innovative Learning Models & Entrepreneurship" were most prominent. China, the UK, and the US were leading contributors. Topic trajectories show rising emphasis on AI-enabled pedagogy and digital integration.
Discussion: The dual bibliometric–semantic approach reveals underexplored themes and actionable directions for digital entrepreneurship and AI-enhanced education, informing policy and institutional strategy.

KEYWORDS

artificial intelligence, entrepreneurship, online education, student, learning

1 Introduction

Artificial intelligence (AI) has transitioned from a niche technological concept into a transformative force across multiple sectors—including healthcare, finance, transportation, and notably, education. In recent years, the integration of AI into online learning environments has accelerated, offering adaptive systems, personalized content delivery, and intelligent tutoring systems that enhance the digital learning experience (Dogan et al., 2023; Willis, 2024). This transformation has not only changed how knowledge is delivered but has also redefined learner engagement, assessment strategies, and the scalability of education.

Simultaneously, entrepreneurship education has gained prominence as a key competence for driving innovation and socio-economic growth in the digital era (Sarumi, 2024; Giuggioli and Pellegrini, 2023). As economies become increasingly knowledge-driven and innovation-led, the cultivation of entrepreneurial mindsets and skills is recognized as essential for preparing individuals to thrive in uncertain, fast-evolving environments. This has led to growing interest in integrating entrepreneurial education into digital learning platforms.

However, existing studies have often explored these domains in isolation—either focusing on pedagogical innovations in AI-enhanced learning or the role of AI in business and entrepreneurship (Bhatia et al., 2024). There remains a significant gap in understanding how AI applications intersect with entrepreneurship education within online learning contexts—an intersection that is critical for equipping learners with future-ready competencies and fostering innovation in the Fourth Industrial Revolution (Darnell and Gopalkrishnan, 2024; Dinger et al., 2024). Bridging this gap is particularly urgent given the rise of digital entrepreneurship and the increasing reliance on technology-driven learning ecosystems to deliver complex, interdisciplinary skill sets.

This study addresses that gap by conducting a comprehensive bibliometric and topic modeling analysis of research on AI applications in entrepreneurship and online education, spanning 2010–2024. The contributions of this study are threefold: (1) it maps trends in publication output and influence, (2) it identifies intellectual and thematic structures in the field through semantic clustering, and (3) it uncovers emerging interdisciplinary synergies that can guide future research and educational practice. Unlike previous works, this study combines bibliometric techniques with semantic modeling to provide a richer understanding of the field's development over time and across disciplines.

To structure this investigation, the study is guided by the following research questions:

> *RQ1:* What are the publication and citation trends of research on AI applications in entrepreneurship and online education?
>
> *RQ2:* Which countries, institutions, and journals are most influential in this research domain?
>
> *RQ3:* What are the major themes and research clusters emerging from bibliometric and topic modeling analyses?
>
> *RQ4:* How have these themes evolved over time, and what directions do they suggest for future research?

By answering these questions, the study aims to illuminate the evolution of this interdisciplinary domain and provide guidance for future scholarships at the intersection of AI, entrepreneurship, and digital education.

2 Literature review: AI-based intelligence applications in entrepreneurship and online education

Artificial Intelligence (AI) has significantly reshaped online education by enabling personalized learning, adaptive feedback, intelligent tutoring systems, and real-time analytics. AI-driven platforms can track learners' behaviors, predict performance, and dynamically adjust instructional content to maximize engagement and success. For example, Dogan et al. (2023) conducted a systematic review of 276 studies and identified three dominant applications of AI in online learning: adaptive and personalized education, behavioral prediction, and algorithmic decision-making. Similarly, Willis (2024) emphasized AI's role in tailoring educational experiences to individual student needs, improving both engagement and retention in digital learning environments.

Parallel to this, AI has also transformed entrepreneurship by enhancing decision-making, streamlining operations, and enabling predictive market analysis. Giuggioli and Pellegrini (2023) proposed the "AI-enabled entrepreneurial process" model, identifying how AI improves opportunity recognition, strategic choices, and business performance across different stages of venture development. Abuzaid and Alsbou (2024) further elaborated how startups leverage AI to gain operational efficiency, navigate funding challenges, and innovate business models through AI-powered decision support systems. Additionally, Usman et al. (2024) reviewed AI applications in customer analytics, product development, and marketing, highlighting both the benefits and ethical considerations of AI integration in entrepreneurial ventures.

While the literature on AI in education and entrepreneurship is growing, studies that examine their intersection remain relatively scarce. However, this niche is rapidly expanding. Darnell and Gopalkrishnan (2024) proposed a pedagogical framework that integrates generative AI into entrepreneurship education, suggesting that AI can support ideation, pitch development, and market analysis as active learning components. Similarly, Dinger et al. (2024) demonstrated the value of AI-driven tools in project-based digital entrepreneurship curricula, emphasizing their role in lowering barriers to entry for novice entrepreneurs.

These insights reflect a shift toward more integrated approaches where AI not only facilitates learning but also acts as a co-creator in entrepreneurial ecosystems, paving the way for the next generation of digitally fluent innovators.

3 Methods

This study employed a bibliometric analysis using the Bibliometrix package to explore the research landscape surrounding artificial intelligence applications in entrepreneurship and online education. The bibliometric analysis identified the most cited countries, key journals contributing to the domain, and prevailing trends. Furthermore, thematic mapping and thematic evolution analyses were conducted to visualize and track the development of key themes over time, providing a comprehensive understanding of the intellectual structure of the field (Aria and Cuccurullo, 2017; Aytekin et al., 2024). This combined approach was chosen to provide both quantitative and semantic insights. Bibliometric analysis offers a macro-level view of scientific activity, while topic modeling uncovers deeper, latent themes from large textual datasets, especially suited for interdisciplinary domains like AI in education and entrepreneurship.

In addition to bibliometric analysis, a topic modeling approach was utilized to identify latent semantic patterns within extensive text datasets. Topic modeling, a probabilistic technique, is instrumental in discovering hidden topics within unstructured documents, capturing the semantic structure of textual data. The methodology is based on the principle that certain terms are more frequent within specific documents due to their relevance to distinct topics (Blei et al., 2003). By examining word co-occurrence patterns, topic modeling unveils these underlying semantic clusters. The analysis involves calculating the probability of distribution of topics across documents, topic distributions within individual documents, and word-level topic assignments (Blei, 2012). This method is particularly appropriate here due to the large volume of interdisciplinary textual data and the need to extract latent thematic structures without supervision. Furthermore, the integration of bibliometric techniques with topic modeling has been increasingly employed to uncover research trends, thematic evolution, and emerging domains in scientific literature (Özköse et al., 2023; Ozyurt and Ayaz, 2022; Ozyurt et al., 2024; Gurcan et al., 2023).

Numerous algorithms have been developed for topic modeling in text mining and natural language processing, including Latent Dirichlet Allocation (LDA), Hierarchical Latent Dirichlet Allocation (HLDA), Hierarchical Dirichlet Process (HDP), Non-Negative Matrix Factorization (NMF), Dirichlet Multinomial Regression (DMR), Dynamic Topic Model (DTM), and Correlated Topic Model (CTM). Models like NMF, CTM, and DMR often face challenges in determining the optimal number of topics through traditional consistency metrics. In contrast, newer approaches, such as HDP and HLDA, offer automated mechanisms for identifying the ideal topic count (Vayansky and Kumar, 2020). However, LDA provides flexibility by allowing manual adjustment of topic numbers, which facilitates iterative fine-tuning for improved accuracy and semantic coherence (Gurcan et al., 2023). This adaptability, combined with effective coherence scoring methods, establishes LDA as a prominent tool across various research domains for analyzing semantic content in large text corpora. Consequently, LDA remains a widely utilized approach in academic and applied research (Blei et al., 2003). Given its adaptability and interpretability, LDA was chosen as the most appropriate model for the current study. A mathematical and graphical representation of the LDA model is illustrated in Figure 1.

Figure 1 illustrates the probabilistic graphical model of the Latent Dirichlet Allocation (LDA) used in our analysis. It depicts the relationships between documents, topics, and words, providing mathematical context for the topic modeling process applied in this study. LDA analyzes the semantic content of documents by exploring latent semantic structures (Griffiths and Steyvers, 2004). The method assigns words in documents to random variables and clusters them based on a probabilistic process driven by the Dirichlet distribution (Blei, 2012). As an unsupervised learning method, LDA eliminates the need for labeled data or training sets, enabling efficient analysis of extensive document collections to extract semantic patterns (Blei et al., 2003).

3.1 Search strategy and study selection

To ensure methodological rigor and comprehensive coverage, this study employed both Scopus and Web of Science (WoS) as primary data sources. Scopus, the largest abstract and citation database of peer-reviewed literature, is widely recognized for its broad coverage of journals, books, and conference proceedings, while WoS is known for indexing high-impact and prestigious academic journals. The combined use of these databases minimized potential bias that could arise from relying on a single source and is consistent with prior bibliometric studies emphasizing the importance of multi-database searches for reliable literature mapping (Falagas et al., 2008).

The time frame of 2010–2024 was deliberately chosen. Earlier works prior to 2010 predominantly addressed theoretical perspectives of artificial intelligence. In contrast, the period after 2010 marks a significant increase in practical implementations of

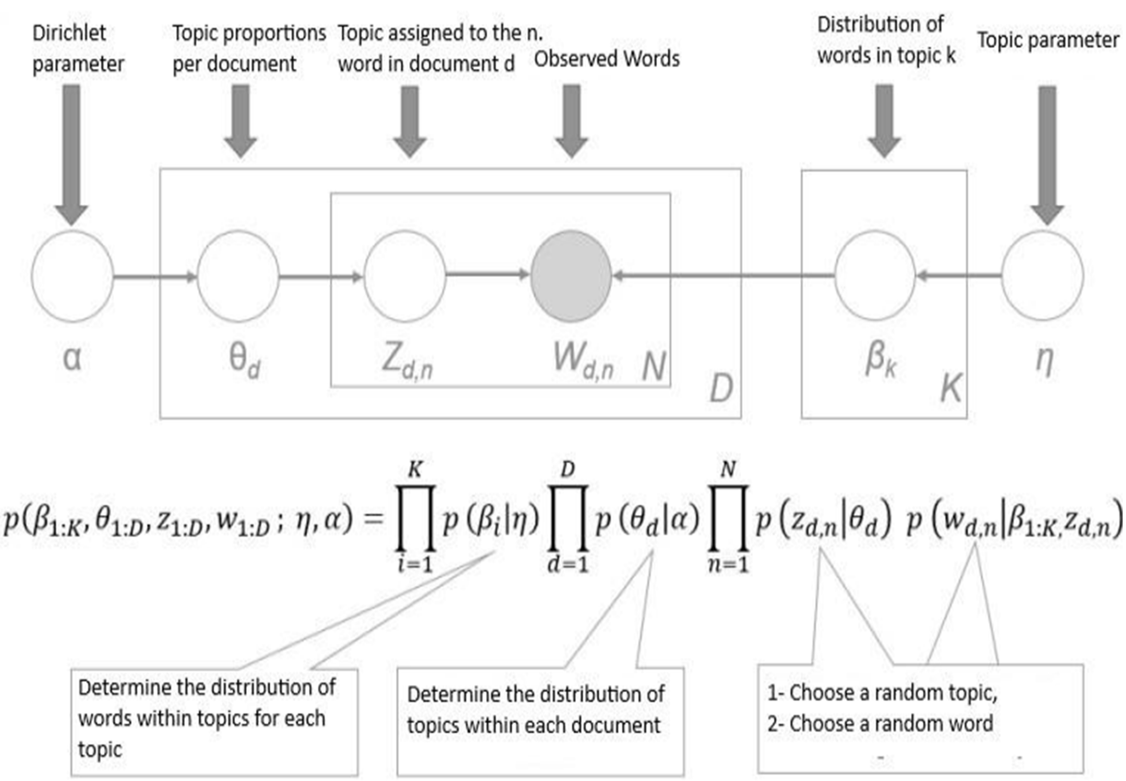

FIGURE 1
LDA mathematical and graphical representation.

AI in education and entrepreneurship, which aligns with the focus of this study on contemporary applications. Restricting the dataset to this time span therefore ensured that the analysis captured the most relevant and impactful developments.

The search query was carefully designed to encompass a wide range of related concepts, combining terms on artificial intelligence, entrepreneurship, and online education. The following Boolean string was applied to both databases, targeting the title, abstract, and keywords of articles published in English:

> ("artificial intelligence" OR "machine learning" OR "deep learning" OR "neural networks" OR "AI-powered" OR "intelligent systems" OR "automated learning" OR "algorithmic learning") AND ("entrepreneurship" OR "innovation" OR "startups" OR "business development" OR "technological advancement" OR "venture creation" OR "business innovation") AND ("online learning" OR "e-learning" OR "elearning" OR "distance education" OR "online education" OR "virtual learning" OR "digital education" OR "remote learning" OR "distance learning" OR "web-based learning" OR "cyber education" OR "blended learning" OR "MOOCs" OR "online courses" OR "digital learning platforms" OR "educational technology" OR "learning management systems" OR "virtual classrooms" OR "computer-assisted instruction" OR "internet-based learning" OR "mobile learning" OR "ubiquitous learning" OR "flipped classroom" OR "synchronous learning" OR "asynchronous learning")

Only peer-reviewed research articles and review articles were included to ensure high-quality scholarly contributions.

The article selection process followed a structured four-step procedure:

1 Initial retrieval of 1,012 records from Scopus and WOS.
2 Deduplication, which resulted in the removal of 123 duplicate records.
3 Screening of titles and abstracts, leading to the exclusion of 400 irrelevant documents.
4 Final inclusion of 489 articles, which formed the dataset for bibliometric and topic modeling analyses.

This transparent and systematic process, combining a robust multi-database search, a justified time frame, and a well-defined selection procedure, provides a reliable foundation for analyzing research trends and thematic evolution in AI applications in entrepreneurship and online education.

3.2 Pre-processing

Preprocessing is a critical step in ensuring the consistency and quality of datasets for analysis. The collected data underwent several cleaning and transformation steps using Python's Natural Language Toolkit (NLTK) (Bird et al., 2010). Texts were standardized by converting all content to lowercase, and irrelevant metadata, such as web links and publisher information, was removed. Tokenization was performed to break down the text into individual words, followed by the removal of English stop words (e.g., "and," "is," "or"), numerical expressions, punctuation, and symbols. Common academic terms like "article," "research," and "study" were excluded to avoid noise in the semantic analysis (Gurcan et al., 2023). Lemmatization was applied to derive meaningful base words, enhancing the semantic coherence of the corpus (Plisson et al., 2004).

An N-gram model was used to identify high-frequency terms, focusing on unigrams to analyze co-occurrence patterns. Each article was subsequently transformed into a word vector using the "bag of words" method, which enabled the creation of a document-term matrix (DTM) (Blei, 2012). This matrix served as the foundation for statistical modeling to extract latent topics.

3.3 Data analysis and fitting topic modeling

The Gensim library in Python facilitated the implementation of the LDA algorithm (Řehůřek and Sojka, 2010). Parameter optimization involved setting α and β to "Symmetric," assuming equal probability distributions of topics across documents and words within topics. The model fitting process employed an iterative and heuristic approach (Ozyurt et al., 2024), with 100 iterations and 15 passes to refine the topic distributions. To improve the quality of the analysis, words with a minimum probability below 0.01 and infrequently occurring terms within the text were excluded from the dataset. Additionally, a predefined list of stop words, commonly used but semantically unimportant terms, was removed to focus on the meaningful content of the documents.

To determine the optimal number of topics (K), models were generated for K values ranging from 5 to 15, and coherence scores were calculated. A coherence score approximating 0.7 is generally considered ideal (Blei et al., 2003). However, in this study, the highest coherence score obtained was 0.3962, leading to the selection of a model with five topics to best balance interpretability and data representation. Although coherence scores remained moderate, the selected model with 5 topics offered the clearest interpretability, justifying its use for mapping thematic clusters in this study.

The LDA model assigned probability to topics within documents and terms within topics. These probabilities informed the ranking of terms and the labeling of topics. The Python pyLDAvis library (Mabey, n.d.; Sik et al., 2023) was employed for topic visualization, aiding the interpretation of results. Researchers, with input from domain experts, labeled topics based on the most representative terms. Additionally, the percentage of each topic within documents, term distributions, and overall topic prevalence across articles were calculated. The top 10 terms with the highest frequencies for each topic were identified.

Temporal changes in topic prevalence were analyzed using percentage change and acceleration metrics calculated in Microsoft Excel. The slope of trends over time was determined using the SLOPE function, where x-values represented years and y-values denoted the frequency of terms or topics. Positive or negative acceleration values indicated an increase or decrease in the rate of publication on specific topics. Graphical visualizations illustrate topic volumes and trends, providing insights into the temporal dynamics of the analyzed research field.

4 Results and discussion

4.1 Bibliometric analysis results and discussion

Figure 2 summarizes the academic growth and research trends of AI applications in the context of entrepreneurship and online education between 2010 and 2024.

This Figure 2, covering the years 2010–2024, summarizes the development of AI applications in the context of entrepreneurship and online education and their growth in academic literature. There are 489 documents and 312 sources in total, and the annual growth rate of the documents is quite impressive at 46.56%, reflecting a pattern consistent with broader trends in AI in education and entrepreneurship research over the past 5 years (Arianti et al., 2024). During this period, 1,563 different authors contributed, and the number of single-authored documents was limited to 82. However, the average number of co-authors per author was determined as 3.64, indicating that a collaborative research environment prevails. In addition, only 6.339% of the documents were written with international collaboration, which can be interpreted as an area where global participation can be increased. The average number of citations per document was calculated as 11.35, which reveals that the academic impact level of the field is high. The number of keywords of the authors is quite diverse with 1,728, indicating that it covers a wide range of topics. In addition, the total number of references is recorded as 16,872, which emphasizes that the studies are based on a comprehensive literature base. The average age of the documents is 2.64 years, indicating that the information and concepts in the studies are up to date. This indicates that literature is constantly renewed in line with the rapidly developing technology and digitalization processes. Overall, these data provide a valuable framework for understanding research trends and collaboration models in entrepreneurship and online education in artificial intelligence applications. Figure 3 shows the distribution of documents published about artificial intelligence and online education by source journals.

The journal Applied Mathematics and Nonlinear Sciences stands out as the journal that publishes the most documents in this field with 26 documents. It is followed by IEEE Access with 20 documents. Sustainability (Switzerland) ranks third with nine documents and emphasizes the intersection of artificial intelligence and educational technologies in the context of sustainability.

Then, journals such as Computer-Aided Design and Applications (eight documents) and Wireless Communications and Mobile Computing (seven documents) attract attention. These journals contribute to the consideration of technological applications in the context of education and entrepreneurship. Other prominent sources include Computational Intelligence and Neuroscience (six documents), Education and Information Technologies (5 documents), Frontiers in Education (five documents), IEEE Transactions on Consumer Electronics (5 documents), and Information (Switzerland) (five documents).

This Figure 3 shows that AI and online education are a multidisciplinary field. Journals that examine the impact of technological innovations in the context of education (e.g., Education and Information Technologies) and more technically focused journals (e.g., IEEE Access, Wireless Communications and Mobile Computing) together guide the studies in this field. The fact that journals that focus on basic sciences, such as Applied Mathematics and Nonlinear Sciences, are also at the forefront in this field shows the importance of the mathematical modeling and application dimensions of AI technologies. This clearly shows that AI applications in entrepreneurship and online education are a multidisciplinary research area, with strong representation from education, business, computer science, and cognitive science fields (Turmuzi and Tyaningsih, 2025). Researchers can contribute to both theoretical and practical innovations by benefiting from the studies published in these journals. Figure 4 shows the total number of citations by country on AI applications and online education.

In Figure 4, China ranks first with a total of 1,022 citations, reflecting its dominant position in AI in education and entrepreneurship research, consistent with other bibliometric mappings in this domain (Hossein-Mohand et al., 2025). China is followed by the United Kingdom with 479 citations and the United States with 405 citations, respectively. These three countries are making a significant academic impact in AI and online education research. India (310 citations) and Mexico (294 citations) stand out

FIGURE 2
Main information.

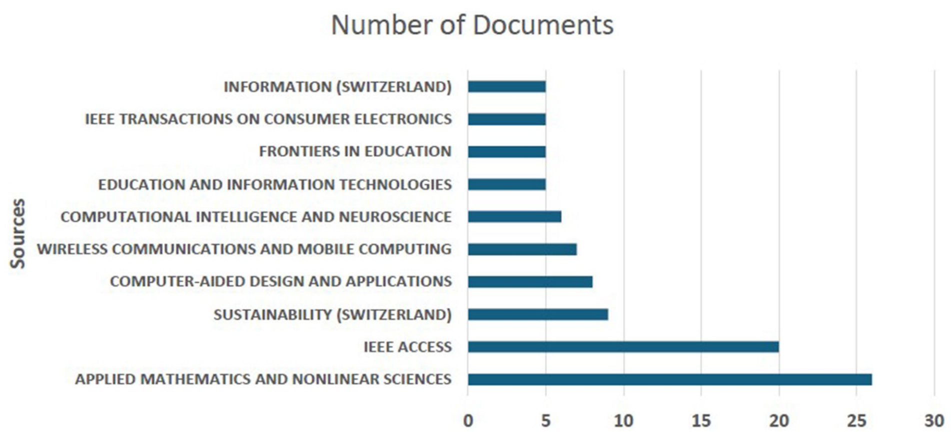

FIGURE 3
The most sources.

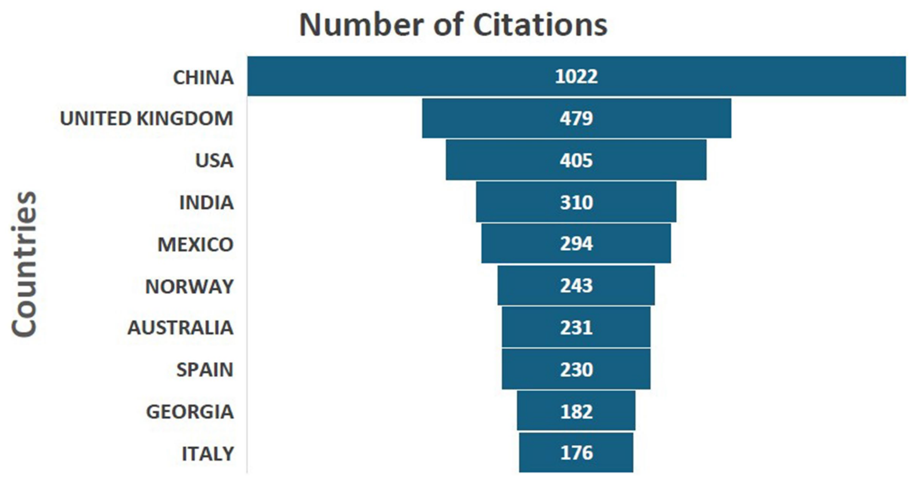

FIGURE 4
The most cited countries.

in the developing countries category and increase their contributions in the field. Countries with fewer citations include Norway (243 citations), Australia (231 citations), and Spain (230 citations). Norway has a high level of academic contribution in AI and online education despite its small population. Georgia (182 citations) and Italy (176 citations) are also on the list with their work in this field. This figure shows that countries such as China, the United Kingdom, and the United States are in academic leadership positions in AI applications. China's leading position shows that it is a leader in both the development of AI technologies and their integration into the education system. The United Kingdom and the United States play an influential role in the application of AI, especially in entrepreneurship education. The growing influence of developing countries, especially India and Mexico, in this area points to the potential for broader academic collaboration. In this context, knowledge transfer between countries and increased international collaboration can contribute to further expanding academic impact and application potential in the field. Citation density reflects both the research capacity of countries and the interest in these studies on a global scale.

The thematic map (Figure 5) shows the analysis of themes in AI and online education in terms of their degree of development (density) and degree of importance (centrality). The themes are divided into four main groups: Basic Themes, Engine Themes, Niche Themes, and Emerging or Declining Themes.

When Figure 5 is examined, it is seen that the themes in artificial intelligence and online education are divided into four main groups in terms of development level (density) and importance level (centrality). Among the motor themes, "Artificial Intelligence," "Higher Education" and "E-learning" attract attention with their high centrality and density values. This situation shows that artificial intelligence and online education are a strategic focus in entrepreneurship and education applications. Among the basic themes, "Machine Learning," "Deep Learning" and "Online Learning" have high centrality but lower density, which reveals that these areas have a wide application potential but need to be developed further.

In addition, "Adaptive Learning System" and "Active Learning," which are considered niche themes, show that despite their high density, they provide significant developments in certain sub-areas with their low

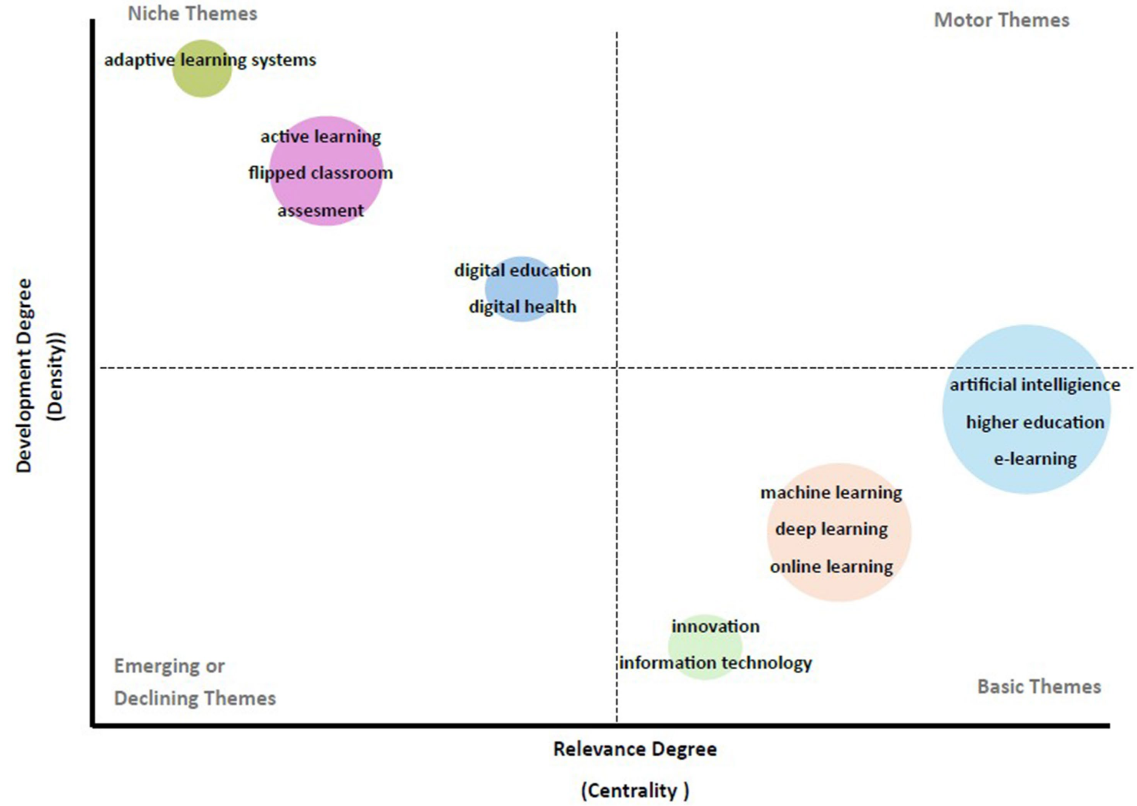

FIGURE 5
Thematic map.

centrality. It can be said that these themes can contribute to individualized and customized learning experiences, especially in entrepreneurship education. The developing or out-of-favor themes, "Digital Education," "Digital Health," "Innovation" and "Information Technology," indicate that they either have development potential or have lost their impact due to their low centrality. It is understood that the themes of digital education and digital health may have gained importance with the COVID-19 pandemic, but their long-term effects should be evaluated.

This analysis highlights the basic and advanced themes in the field of artificial intelligence and online education, as well as niche and emerging areas that promise potential. While engine themes provide a strategic focus, niche themes offer opportunities for developing customized solutions. In this context, the current thematic structure of artificial intelligence applications in entrepreneurship and online education sheds light on the future development of the field. Figure 6 illustrates the thematic evolution between 2010–2019 and 2020–2024.

When Figure 6 is examined, it is clearly seen how thematic areas in the 2010–2019 period have evolved into the 2020–2024 period. It is noteworthy that themes such as "Machine Learning" and "Artificial Intelligence" have maintained their importance in the 2020–2024 period, and the "Artificial Intelligence" theme has come to the forefront more clearly. This shows that artificial intelligence plays a central role in both education and entrepreneurship applications. In addition, it is understood that the "Higher Education" theme has continued and the integration of technology into the education system is increasingly emphasized. The "Adaptive Learning System" theme continues to exist in both periods, indicating that personalized learning systems maintain their importance. The "Blended Learning" theme that emerged in the 2020–2024 period shows that the popularity of hybrid education models has increased, and that online learning and face-to-face education are used in a balanced way (Redondo-Rodríguez et al., 2025). The theme of "Technology Adoption" continues to exist in both periods, revealing that the integration of technology into education and entrepreneurship processes is a critical issue. However, the theme of "Internet of Things," which attracted attention in the 2010–2019 period, seems to have lost its importance in the 2020–2024 period. This may indicate that the focus has shifted to more specific technologies. This thematic evolution indicates that there is an increasing focus on the applications of artificial intelligence and technology in education systems and entrepreneurship, especially post-pandemic, as seen in recent bibliometric trends (Jantakun et al., 2024). Artificial intelligence is expected to be adopted more in areas such as higher education and hybrid learning models. Figure 7 demonstrates the trend analysis of key terms between 2018 and 2024.

Terms such as "Artificial Intelligence," "Blended Learning," and "Deep Learning" have shown consistent growth, particularly gaining momentum after 2020. The term "COVID-19" emerges during 2020, highlighting its influence on online learning and digital education. Additionally, terms like "Higher Education" and "Educational Technology" remain prominent throughout the period, reflecting their critical role in advancing technological adoption in education. Meanwhile, earlier terms such as "Flipped Classroom" and "Active Learning" show diminishing relevance in

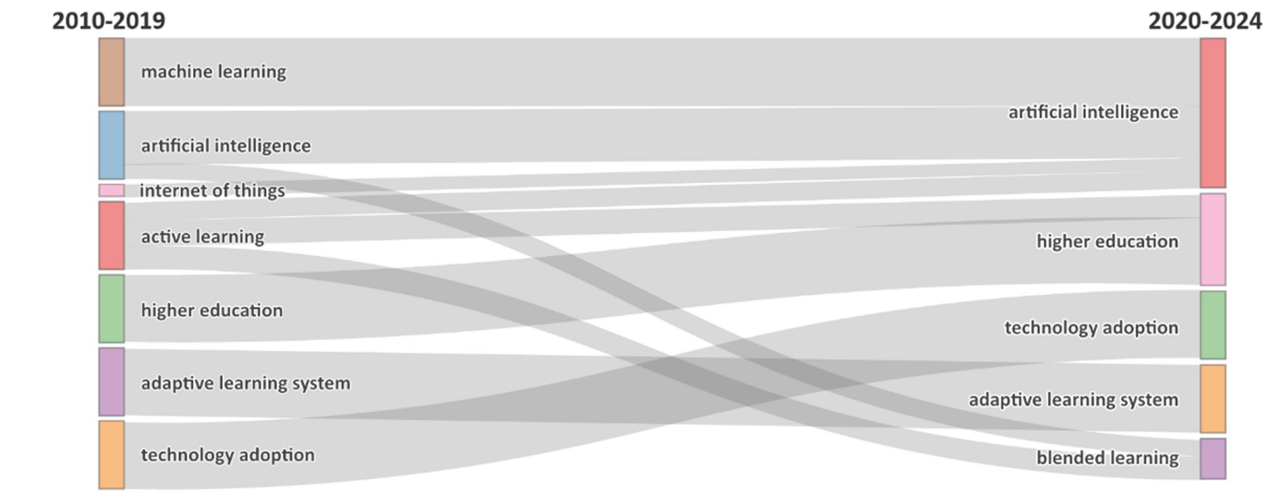

FIGURE 6
Thematic evolution.

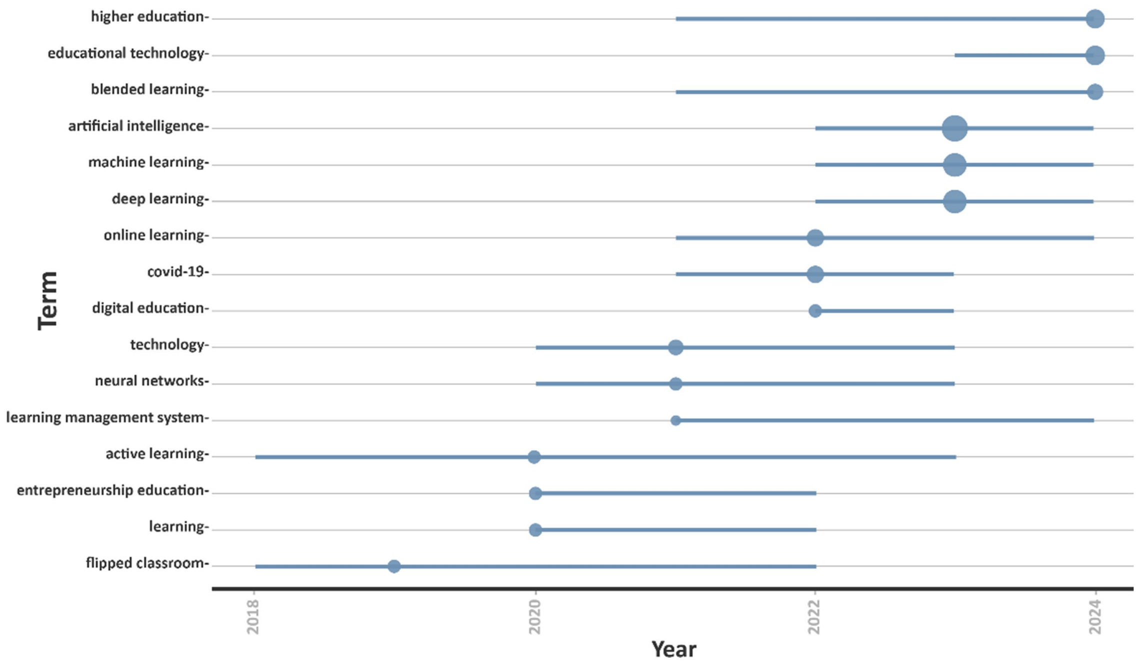

FIGURE 7
Trend topics.

recent years. This trend analysis illustrates the evolving focus toward AI-driven and blended educational methodologies in response to emerging challenges and innovations.

4.2 Topic modeling results and discussion

This study employs topic modeling to examine prominent themes, a method widely used in AI-education bibliometric studies to detect latent semantic structures (Blanco-González-Tejero et al., 2023). The results of the topic modeling analysis are presented alongside the distribution and temporal trends of the identified topics. By highlighting topics with increasing or decreasing relevance, the analysis underscores the dynamic nature of the research field.

The findings provide valuable insights into which topics have gained traction, and which have diminished in significance, offering a comprehensive perspective on the development of the research landscape. This approach not only identifies emerging areas of interest but also pinpoints those becoming less central, thereby enhancing the understanding of the field's progression and potential future directions.

Using Latent Dirichlet Allocation (LDA) analysis on the selected publications, five distinct topics were identified, each characterized by its top 10 associated keywords. Table 1 presents a summary of the topic labels, associated terms, and their respective percentages. The keywords define the core themes of each topic, while the percentages represent the proportion of the corpus covered by each topic. This enables a clearer comprehension of the dominant research areas and the relative significance of each topic within the dataset.

Upon examining Table 1, five key topics emerge as focal points within the dataset. The most prominent topic, "Teaching and Pedagogical Innovations" (24%), highlights advancements in teaching practices and pedagogical approaches. Terms such as "teaching" (4.10%), "education" (3.75%), and "student" (3.15%) reflect the centrality of improving educational methods and integrating innovative technologies. This underscores the focus on student-centered approaches and modern pedagogical frameworks.

The second topic, "Innovative Learning Models and Entrepreneurship" (23%), emphasizes the development of novel learning frameworks and their intersection with entrepreneurial activities. Key terms such as "learning" (5.23%), "model" (2.91%), and "online" (1.57%) suggest a growing emphasis on virtual learning environments, while terms like "deep" (1.72%), "machine" (1.12%), and "algorithm" (1.06%) indicate the application of advanced computational methods in education.

The third topic, "Educational Technology and Digital Integration" (20%), underscores the increasing role of digital tools and artificial intelligence in educational practices. Terms such as "learning" (3.74%), "technology" (2.65%), and "artificial intelligence" (1.89%) reveal how digital applications are transforming teaching and learning processes. This topic aligns with the broader trend of leveraging technology to enhance educational outcomes.

"Technological Advances in Health and Education" (19%), the fourth topic, bridges the domains of health and education by emphasizing technological innovations. Terms like "learning" (4.41%), "health" (1.85%), and "virtual" (1.21%) suggest the integration of virtual and mobile technologies within these interdisciplinary fields, highlighting the convergence of health and educational advancements.

Finally, "Digital, Medical, and Industrial Applications in Education" (14%) focuses on the intersection of education with digital, medical, and industrial sectors. The prominence of terms such as "digital" (7.12%), "medical" (1.81%), and "industry" (1.40%) reflects the growing importance of digital transformation and its implications for interdisciplinary educational applications.

Table 2 illustrates the distribution of topics across different time periods, highlighting the shifting focus of academic inquiry. This analysis not only captures the historical trajectory of research themes but also offers valuable insights into the development and future directions of the field.

As shown in Table 2, the years from 2010 to 2024 were divided into three periods, allowing for an analysis of topic evolution over time. The findings highlight significant shifts in academic research focus, with emerging trends gaining momentum in recent years.

During the first period (2010–2014), research activity was limited across most topics, with only 10 studies in total. This period reflects an early stage of exploration, where interest in Innovative Learning Models and Entrepreneurship and Technological Advances in Health and Education began to take shape, contributing four studies each.

In the second period (2015–2019), research activity increased significantly, with a total of 33 studies across topics. The most notable growth was observed in Teaching and Pedagogical Innovations (12 studies) and Technological Advances in Health and Education (11 studies), indicating a growing emphasis on improving teaching methods and integrating technological advancements into education. This period also marked the emergence of Educational Technology and Digital Integration as a research focus, albeit with a modest increase of three studies.

TABLE 1 Topic distribution and key terms.

PyLDAvis	Topics	Terms with percentages	Percentage
2	Teaching and Pedagogical Innovations	teaching (4.10%) | education (3.75%) | student (3.15%) | learning (2.86%) | technology (2.08%) | innovation (1.52%) | online (1.51%) | model (1.30%) | artificial (1.12%) | intelligence (1.09%)	24%
1	Innovative Learning Models and Entrepreneurship	learning (5.23%) | model (2.91%) | deep (1.72%) | online (1.57%) | data (1.56%) | network (1.53%) | system (1.51%) | student (1.41%) | machine (1.12%) | algorithm (1.06%)	23%
3	Educational Technology and Digital Integration	learning (3.74%) | education (3.28%) | technology (2.65%) | educational (2.57%) | intelligence (2.07%) | artificial (1.89%) | student (1.12%) | application (1.02%) | tool (0.94%) | language (0.90%)	20%
4	Technological Advances in Health and Education	learning (4.41%) | health (1.85%) | technology (1.60%) | student (1.21%) | virtual (1.21%) | system (1.07%) | innovation (1.07%) | mobile (0.98%) | intelligence (0.85%) | artificial (0.83%)	19%
5	Digital, Medical, and Industrial Applications in Education	digital (7.12%) | education (3.33%) | medical (1.81%) | learning (1.70%) | technology (1.53%) | data (1.40%) | industry (1.40%) | development (1.35%) | transformation (1.31%) | innovation (1.30%)	14%

TABLE 2 Evolution of research topics.

Topics	2010–2014	2015–2019	2020–2024	Total	Acc
Teaching and Pedagogical Innovations	1	12	123	136	2.56
Innovative Learning Models and Entrepreneurship	4	6	125	135	2.65
Educational Technology and Digital Integration	1	3	89	93	1.94
Technological Advances in Health and Education	4	11	66	81	1.36
Digital, Medical, and Industrial Applications in Education	0	1	43	44	0.85
Total	10	33	446	489	

The third period (2020–2024) represents a transformative phase in the AI-entrepreneurship-education nexus, marked by the mainstreaming of tools like ChatGPT and intelligent learning environments (Mojolou et al., 2024). The rapid growth in this period is primarily driven by Teaching and Pedagogical Innovations (123 studies) and Innovative Learning Models and Entrepreneurship (125 studies), signaling a heightened interest in innovative approaches to education. Furthermore, Educational Technology and Digital Integration (89 studies) and Technological Advances in Health and Education (66 studies) also gained significant traction, reflecting the increasing reliance on digital tools and health-focused educational methods. Digital, Medical, and Industrial Applications in Education, although representing a smaller share (43 studies), also experienced growth, underscoring its emerging relevance.

The acceleration rates (Acc) provide further insights into the momentum of these topics. Innovative Learning Models and Entrepreneurship has the highest acceleration rate (2.65), followed closely by Teaching and Pedagogical Innovations (2.56). These rates indicate these topics' rapid growth and centrality in recent research. In contrast, Digital, Medical, and Industrial Applications in Education has the lowest acceleration rate (0.85), suggesting its niche yet growing status within the broader research landscape.

Table 3 provides an analysis of the slopes of each topic, both within their individual trajectories and relative to other topics. This dual perspective highlights the internal growth trends of each topic over time while offering a comparative view of their progression in relation to other research areas. Such an approach enables a comprehensive understanding of how specific topics have evolved independently and in the broader context of the research field, thereby identifying not only areas of rapid growth but also those maintaining a steady or declining trajectory relative to their counterparts.

When analyzing Table 3, Digital, Medical, and Industrial Applications in Education exhibits the highest slope in its own development (48.86), suggesting rapid growth and increasing prominence within the research field. Educational Technology and Digital Integration follows closely with a slope of 47.31, reflecting the consistent integration of digital tools and educational technologies over time.

Topics such as Teaching and Pedagogical Innovations (44.85) and Innovative Learning Models and Entrepreneurship (44.81) also demonstrate significant growth, highlighting the field's focus on innovative and pedagogical advancements. In contrast, Technological Advances in Health and Education shows the lowest slope (38.27), indicating a slower rate of development compared to other topics.

When comparing the momentum of topics relative to others over time, Teaching and Pedagogical Innovations leads with the highest acceleration (8.79), showing that it has gained more prominence in comparison to other topics. Educational Technology and Digital Integration (4.98) and Digital, Medical, and Industrial Applications in Education (4.82) also exhibit positive momentum, reflecting their increasing importance in the academic landscape. However, Innovative Learning Models and Entrepreneurship (−5.99) and Technological Advances in Health and Education (−12.60) display negative momentum values, suggesting a relative decline in interest or prioritization compared to other topics.

This dual analysis of within-topic development and comparative momentum provides a nuanced understanding of the dynamics within the research field. Topics like Digital, Medical, and Industrial Applications in Education and Educational Technology and Digital Integration demonstrate rapid growth and increasing relevance, both individually and in comparison, to others. In contrast, the relative decline in momentum for Innovative Learning Models and Entrepreneurship and Technological Advances in Health and Education may indicate shifting priorities or saturation within these areas. These findings highlight the evolving focus of academic research, revealing both emerging areas of interest and those that may require renewed attention to sustain their impact.

The PyLDAvis tool was employed to visualize the identified topics, providing an interactive, web-based interface to explore the topic modeling results. PyLDAvis facilitates the interpretation of topics by representing their distributions on a two-dimensional plane. The ranking of words within topics was determined using a λ value of 0.6, ensuring a balanced representation of term relevance and specificity.

In the visualization, topics are depicted as bubbles on the left panel, where the size of each bubble corresponds to the topic's overall prevalence in the dataset. The proximity between the bubbles reflects the semantic similarity of the topics, with closer bubbles indicating higher overlap in shared words, while more distant bubbles suggest less similarity.

The right-hand panel complements this by displaying the frequency of words. The blue bars represent the overall frequency of a word across the entire dataset, while the red bars indicate the word's specific contribution to a particular topic. This dual representation enables a more detailed exploration of the key terms defining each topic.

Figure 8 provides a screenshot of the PyLDAvis visualization, showcasing semantic relationships and the distribution of topics within the analyzed corpus. This tool allows for a deeper understanding of the latent structures within the data, making it a valuable resource for interpreting topic modeling results.

Figure 8 provides a comprehensive visualization of the topic modeling results using PyLDAvis. The left panel, the Intertopic Distance Map, displays the semantic relationships between topics, where each

TABLE 3 Development of topics over time.

Development of topics over time		Interrelations between topics over time	
Topics	Acc	Topics	Acc
Digital, Medical, and Industrial Applications in Education	48.86	Teaching and Pedagogical Innovations	8.79
Educational Technology and Digital Integration	47.31	Educational Technology and Digital Integration	4.98
Teaching and Pedagogical Innovations	44.85	Digital, Medical, and Industrial Applications in Education	4.82
Innovative Learning Models and Entrepreneurship	44.81	Innovative Learning Models and Entrepreneurship	−5.99
Technological Advances in Health and Education	38.27	Technological Advances in Health and Education	−12.60

circle represents a topic, and its size corresponds to its prevalence in the dataset. Topics closer together share more semantic overlaps, while those farther apart indicate distinct thematic content. For instance, Topic 4 (highlighted in red) occupies a central position, suggesting moderate semantic connections with other topics, whereas Topic 5 is positioned farther away, indicating less overlap with Topic 4 and other topics.

The right panel presents the Top 30 Most Relevant Terms for the selected topic (Topic 4) using two bars for each term. The blue bar represents the overall frequency of the term across all topics, while the red bar indicates its frequency specifically within Topic 4. Terms such as "health," "learning," "mobile," and "virtual" dominate Topic 4, with "health" being particularly significant. This highlights the thematic focus of Topic 4 on healthcare, learning, and the integration of mobile and virtual technologies. The visualization effectively illustrates the semantic structure of the topics, emphasizing the connections and distinctiveness between them. Topic 4's proximity to Topics 2 and 3 suggests shared themes, while its distance from Topic 5 reflects its unique focus. This representation enables an in-depth understanding of the relationships between topics and their defining terms, offering valuable insights into the dataset's thematic composition.

5 Conclusion and recommendations

5.1 Summary of key findings

This study examined 489 documents published between 2010 and 2024 on artificial intelligence (AI) applications in entrepreneurship and online education. The results indicate a rapid growth in the field, with an annual publication rate of 46.56% and a strong citation impact (11.35 citations per document on average). China, the United Kingdom, and the United States emerged as leading contributors, while countries such as India and Mexico showed growing influence. Thematic mapping revealed Artificial Intelligence, Higher Education, and E-learning as motor themes, while topic modeling highlighted five dominant clusters: pedagogical innovations, innovative learning models and entrepreneurship, digital integration, technological advances in health and education, and interdisciplinary digital/industrial applications.

5.2 Theoretical implications

This study advances theory at the intersection of AI-enhanced education and entrepreneurship by demonstrating how AI-supported pedagogy, including adaptive and blended learning models as well as intelligent tutoring systems, co-evolves with data-driven entrepreneurial learning and broader processes of digital integration. The combined bibliometric and LDA approach reveals that core constructs such as machine and deep learning and adaptive systems have migrated from peripheral to central themes over time, while entrepreneurial competences such as creativity, opportunity recognition, and data-driven decision making are increasingly embedded within online learning ecosystems. Moreover, the cross-sectoral linkages that emerge in areas like health and industrial education extend the boundaries of technology-adoption and digital-innovation theories. By aligning topic trajectories with citation structures, our results provide a mechanism to reconcile scientific

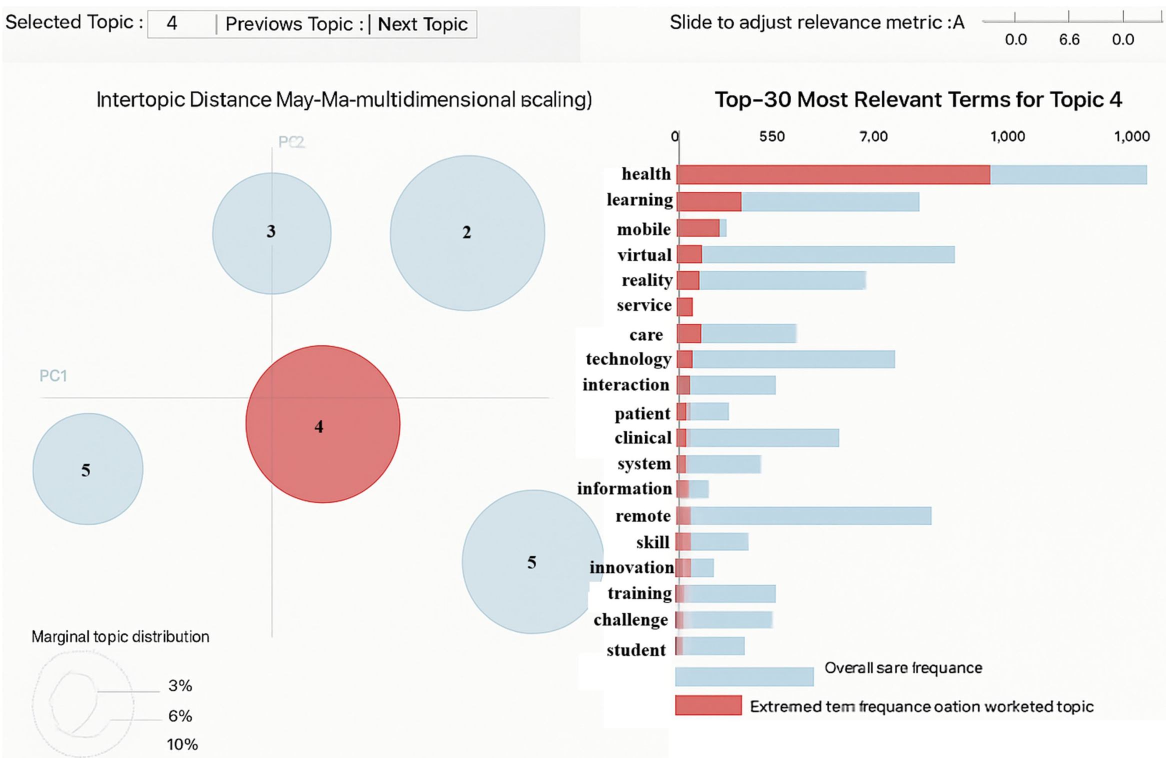

FIGURE 8
PyLDAvis.

impact with semantic novelty—an integration that recent studies in ScienceDirect, SpringerLink, Emerald, and Taylor & Francis have also recommended for tracking conceptual change through the fusion of bibliometric mapping and topic modeling.

5.3 Practical implications

From a practical perspective, the findings suggest that educators can benefit from translating the clusters of "Teaching and Pedagogical Innovations" and "Educational Technology and Digital Integration" into adaptive and blended course blueprints that incorporate learning analytics dashboards, mastery-based progression, and AI-assisted feedback loops. For entrepreneurship education, the "Innovative Learning Models and Entrepreneurship" cluster highlights the potential to establish data-driven venture studios within courses, enabling students to conduct market analyses with machine learning, prototype rapidly with generative AI tools, and develop evidence-based pitching skills. Assessment and quality assurance processes can also be restructured to align rubrics with AI-enabled outcomes such as creativity, problem framing, and data literacy, while ensuring traceability of AI tools through model documentation and auditability. At the institutional level, it becomes essential to prioritize AI literacy training for instructors, invest in privacy-preserving data pipelines, and adopt interoperable educational technology infrastructures that can be scaled across different faculties.

5.4 Policy implications

The policy implications of this study point toward the necessity of targeted funding and incentives for AI-ready digital infrastructure, the creation of open educational resources, and the establishment of interdisciplinary laboratories that connect entrepreneurship with health and industrial education. To address the current low levels of international co-authorship, policymakers should design programs that promote both South–South and North–South collaborations, as well as mobility schemes and shared repositories that lower barriers for researchers in emerging regions. Standards and ethics should also be prioritized, with AI literacy outcomes integrated into program accreditation requirements, privacy-by-design frameworks embedded into institutional practices, and transparency ensured through mandatory model documentation and dataset statements. Furthermore, governments and educational authorities can strengthen their monitoring capacity by developing national dashboards that track topic-level momentum—such as the acceleration of adaptive learning compared to the decline of IoT-related research—thereby supporting evidence-based curriculum updates and teacher development.

5.5 Contributions to knowledge

In addition to mapping the rapid growth, thematic evolution, and emerging frontiers of AI applications in entrepreneurship and online education, this study exemplifies how the integration of bibliometric indicators with topic-level semantics provides a more sensitive and nuanced map of research development than either method alone. This contribution is consistent with recent integrative reviews published in outlets such as ScienceDirect and Emerald, which similarly highlight the value of combining citation-based impact with semantic modeling to capture the complexity of interdisciplinary fields.

5.6 Directions for future research

The topic modeling results not only reveal the dominant themes in AI applications within entrepreneurship and online education but also point to specific questions that can guide subsequent research. As summarized in Table 4, future studies may examine how AI-driven pedagogical models improve learning outcomes across disciplines, in what ways AI-based environments can cultivate entrepreneurial skills and innovation, and what the long-term implications of digital integration are for equity and institutional performance. In addition, opportunities exist to explore how virtual and AI-based technologies can enhance health education and professional training, as well as to investigate the challenges and prospects of adopting AI in interdisciplinary contexts such as medical and industrial education. These research questions provide a structured roadmap for advancing both theoretical and applied dimensions of the field.

5.7 Limitations

Despite its contributions, this study has several limitations that should be acknowledged. Although both Scopus and Web of Science were employed to minimize bias, the indexing policies and metadata quality of these databases still shape the results, and relevant studies available in other databases or gray literature may have been overlooked. The corpus was limited to English-language publications, which means that non-English contributions—particularly those from emerging regions—are likely underrepresented. Citation-based indicators also introduce a time-lag bias by favoring older works; although this was mitigated by coupling citations with topic trajectories, it may still undervalue very recent yet potentially high-impact studies. Finally, modeling decisions related to LDA, including preprocessing, stop-word lists, and parameterization, influence topic coherence. In this study, the coherence scores were moderate, and interpretability was prioritized.

TABLE 4 Future research questions based on topic modeling results.

Topic area	Future research questions
Teaching and Pedagogical Innovations	How can AI-driven pedagogical models improve learning outcomes across different disciplines?
Innovative Learning Models & Entrepreneurship	In what ways can AI-based environments foster entrepreneurial skills and innovation?
Educational Technology and Digital Integration	What are the long-term implications of AI-enabled digital tools for equity and institutional performance in education?
Technological Advances in Health and Education	How can virtual and AI-based technologies enhance health education and professional training?
Digital, Medical, and Industrial Applications	What opportunities and challenges arise from adopting AI in interdisciplinary fields such as medical and industrial education?

Future research could address these limitations by applying alternative models such as structural topic modeling, hierarchical or correlated topic models, and by validating topics through multi-label expert panels.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Author contributions

KA: Conceptualization, Investigation, Writing – original draft. AA: Formal analysis, Methodology, Writing – original draft. FA: Resources, Visualization, Writing – review & editing. ZA: Data curation, Project administration, Validation, Writing – review & editing. DSB: Software, Writing – review & editing. GD: Methodology, Supervision, Writing – review & editing.

Funding

The author(s) declare that no financial support was received for the research and/or publication of this article.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The authors declare that no Gen AI was used in the creation of this manuscript.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Abuzaid, A. N., and Alsbou, M. K. K. (2024). "AI and entrepreneurship: enablers, obstacles, and startups' role in shaping the future economy" in Proceedings of the 2024 international conference on knowledge engineering and communication systems (ICKECS) (New York City: IEEE), 1–6.

Aria, M., and Cuccurullo, C. (2017). Bibliometrix: an R-tool for comprehensive science mapping analysis. *J. Informetr.* 11, 959–975. doi: 10.1016/j.joi.2017.08.007

Arianti, J., Oktaviani, M., Najdah, T., and Sinambela, Y. P. P. (2024). Analisis Bibliometrik Publikasi Perkembangan Artificial Intelegent (AI) Dalam Dunia Pendidikan. *AKADEMIK: Jurnal Mahasiswa Humanis* 4, 477–487. doi: 10.37481/jmh.v4i2.854

Aytekin, A., Ayaz, R., and Ayaz, A. (2024). Impact of the COVID-19 pandemic on subjective well-being and quality of life: a comprehensive bibliometric and thematic analysis. *Heal. Care Anal.* doi: 10.1007/s10728-024-00507-w

Bhatia, D. A., Bhatia, P., and Sood, D. (2024). Leveraging AI to transform online higher education: focusing on personalized learning, assessment, and student engagement. *Int. J. Manag. Humanit.* 11, 1–6. doi: 10.35940/ijmh.A1753.11010924

Bird, S., Klein, E., and Loper, E. (2010). Natural language processing with Python: analyzing text with the natural language toolkit. Santa Rosa: O'Reilly.

Blanco-González-Tejero, C., Ribeiro-Navarrete, B., Cano-Marin, E., and McDowell, W. C. (2023). A systematic literature review on the role of artificial intelligence in entrepreneurial activity. *Int. J. Semant. Web Inf. Syst.* 19, 1–16. doi: 10.4018/IJSWIS.318448

Blei, D. M. (2012). Probabilistic topic models. *Commun. ACM* 55, 77–84. doi: 10.1145/2133806.2133826

Blei, D. M., Ng, A. Y., and Jordan, M. I. (2003). Latent dirichlet allocation. *J. Mach. Learn. Res.* 3, 993–1022. doi: 10.1016/b978-0-12-411519-4.00006-9

Darnell, J. A., and Gopalkrishnan, S. (2024). Entrepreneurship teaching exercises: integrating generative AI. *Discov. Educ.* 3:172. doi: 10.1007/s44217-024-00261-0

Dinger, B., Byun, S.-E., and Park, J. (2024). Integrating AI into curricula: project-based learning in digital entrepreneurship. *Int. J. Fash. Des. Technol. Educ.*, 1–10. doi: 10.1080/17543266.2024.2438170

Dogan, M. E., Goru Dogan, T., and Bozkurt, A. (2023). The use of artificial intelligence (AI) in online learning and distance education processes: a systematic review of empirical studies. *Appl. Sci.* 13:3056. doi: 10.3390/app13053056

Falagas, M. E., Pitsouni, E. I., Malietzis, G. A., and Pappas, G. (2008). Comparison of PubMed, Scopus, web of science, and Google scholar: strengths and weaknesses. *FASEB J.* 22, 338–342. doi: 10.1096/fj.07-9492lsf

Giuggioli, G., and Pellegrini, M. M. (2023). Artificial intelligence as an enabler for entrepreneurs: a systematic literature review and an agenda for future research. *Int. J. Entrep. Behav. Res.* 29, 816–837. doi: 10.1108/IJEBR-05-2021-0426

Griffiths, T. L., and Steyvers, M. (2004). Finding scientific topics. *Proc. Natl. Acad. Sci. USA* 101, 5228–5235. doi: 10.1073/pnas.0307752101

Gurcan, F., Ayaz, A., Menekse Dalveren, G. G., and Derawi, M. (2023). Business intelligence strategies, best practices, and latest trends: analysis of scientometric data from 2003 to 2023 using machine learning. *Sustainability* 15:9854. doi: 10.3390/su15139854

Hossein-Mohand, H., Hossein-Mohand, H., Albanese, V., and Olmos Gómez, M. d. C. (2025). AI in mathematics education: a bibliometric analysis of global trends and collaborations (2020-2024). *Eurasia J. Math. Sci. Technol. Educ.* 21:em2576. doi: 10.29333/ejmste/15915

Jantakun, T., Jantakun, K., and Jantakoon, T. (2024). Bibliometric analysis of artificial intelligence in STEM education. *High. Educ. Stud.* 15:69. doi: 10.5539/hes.v15n1p69

Mabey, B. PyLDAvis 3.4.1. Available online at: https://pypi.org/project/pyLDAvis/ (Accessed June 1, 2023).

Mojolou, D. N., Madli, F., Thomas, H., Sondoh Junior, S. L., and Karunia, E. (2024). A trend in publishing artificial intelligence personalized education: enhancing learning experiences for university students. *J. Inf. Syst. Technol. Manag.* 9, 286–301. doi: 10.35631/JISTM.937021

Özköse, H., Ozyurt, O., and Ayaz, A. (2023). Management information systems research: a topic modeling based bibliometric analysis. *J. Comput. Inf. Syst.* 63, 1166–1182. doi: 10.1080/08874417.2022.2132429

Ozyurt, O., and Ayaz, A. (2022). Twenty-five years of education and information technologies: insights from a topic modeling based bibliometric analysis. *Educ. Inf. Technol.* 27, 11025–11054. doi: 10.1007/s10639-022-11071-y

Ozyurt, O., Kabakus, A. K., and Ayaz, A. (2024). "Exploring scholarly journal content: topic modeling analysis using LDA algorithm" in Text mining in educational research (Singapore: Springer Nature Singapore), 71–95.

Ozyurt, O., Özköse, H., and Ayaz, A. (2024). Evaluating the latest trends of industry 4.0 based on LDA topic model. *J. Supercomput.* 80, 19003–19030. doi: 10.1007/s11227-024-06247-x

Plisson, J., Lavrac, N., and Mladenić, D. D. (2004). A rule based approach to word lemmatization. In Proceedings of the 7th international multiconference information society (IS'04), pp. 83–86.

Redondo-Rodríguez, M. D., Díaz-Garrido, E., Pérez-Bustamante Yábar, D. C., and Ramón-Jerónimo, M. Á. (2025). Entrepreneurship and artificial intelligence: a bibliometric analysis. *J. Technol. Transf.* 50, 1840–1872. doi: 10.1007/s10961-024-10165-8

Řehůřek, R., and Sojka, P. (2010) Software Framework for Topic Modelling with Large Corpora. In Proceedings of the LREC 2010 Workshop on New Challenges for NLP Frameworks. *Valletta, Malta: European Language Resources Association (ELRA)*, 45–50.

Sarumi, O. O. (2024). AI: driving entrepreneurship, nurturing innovation, and fueling startup growth. *Int. J. Adv. Eng. Manag.* 6, 540–547. doi: 10.35629/5252-0611540547

Sik, D., Németh, R., and Katona, E. (2023). Topic modelling online depression forums: beyond narratives of self-objectification and self-blaming. *J. Ment. Health* 32, 386–395. doi: 10.1080/09638237.2021.1979493

Turmuzi, M., and Tyaningsih, R. Y. (2025). A bibliometric analysis of the development of artificial intelligence (AI) research in education in Scopus indexed journals: what are the future trends of this research? *TEM J.* 14, 671–683. doi: 10.18421/TEM141-60

Usman, F. O., Eyo-Udo, N. L., Etukudoh, E. A., Odonkor, B., Ibeh, C. V., and Adegbola, A. (2024). A critical review of AI-driven strategies for entrepreneurial success. *Int. J. Manag. Entrep. Res.* 6, 200–215. doi: 10.51594/ijmer.v6i1.748

Vayansky, I., and Kumar, S. A. P. (2020). A review of topic modeling methods. *Inf. Syst.* 94:101582. doi: 10.1016/j.is.2020.101582

Willis, V. (2024). The role of artificial intelligence (AI) in personalizing online learning. *J. Online Distance Learn.* 3, 1–13. doi: 10.47941/jodl.1689