Computers and Education: Artificial Intelligence 1 (2020) 100002

Contents lists available at ScienceDirect

Computers and Education: Artificial Intelligence

journal homepage: www.elsevier.com/locate/caeai

# Application and theory gaps during the rise of Artificial Intelligence in Education

Xieling Chen [a], Haoran Xie [b,*], Di Zou [c], Gwo-Jen Hwang [d]

[a] Department of Mathematics and Information Technology, The Education University of Hong Kong, Hong Kong SAR, 10 Lo Ping Road, Tai Po, New Territories, Hong Kong
[b] Department of Computing and Decision Sciences, Lingnan University, Hong Kong SAR, 8 Castle Peak Road, Tuen Mun, New Territories, Hong Kong
[c] Department of English Language Education, The Education University of Hong Kong, Hong Kong SAR, 10 Lo Ping Road, Tai Po, New Territories, Hong Kong
[d] Graduate Institute of Digital Learning and Education, National Taiwan University of Science and Technology, Taiwan

ARTICLE INFO

*Keywords:*
Artificial intelligence in education
Systematic review
Application gap
Theory gap

ABSTRACT

Considering the increasing importance of Artificial Intelligence in Education (AIEd) and the absence of a comprehensive review on it, this research aims to conduct a comprehensive and systematic review of influential AIEd studies. We analyzed 45 articles in terms of annual distribution, leading journals, institutions, countries/regions, the most frequently used terms, as well as theories and technologies adopted. We also evaluated definitions of AIEd from broad and narrow perspectives and clarified the relationship among AIEd, Educational Data Mining, Computer-Based Education, and Learning Analytics. Results indicated that: 1) there was a continuingly increasing interest in and impact of AIEd research; 2) little work had been conducted to bring deep learning technologies into educational contexts; 3) traditional AI technologies, such as natural language processing were commonly adopted in educational contexts, while more advanced techniques were rarely adopted, 4) there was a lack of studies that both employ AI technologies and engage deeply with educational theories. Findings suggested scholars to 1) seek the potential of applying AI in physical classroom settings; 2) spare efforts to recognize detailed entailment relationships between learners' answers and the desired conceptual understanding within intelligent tutoring systems; 3) pay more attention to the adoption of advanced deep learning algorithms such as generative adversarial network and deep neural network; 4) seek the potential of NLP in promoting precision or personalized education; 5) combine biomedical detection and imaging technologies such as electroencephalogram, and target at issues regarding learners' during the learning process; and 6) closely incorporate the application of AI technologies with educational theories.

## 1. Introduction

Artificial Intelligence in Education (AIEd) concerns mainly about the development of "computers which perform cognitive tasks, usually associated with human minds, particularly learning and problem-solving (p. 10)" (Baker and Smith (2019). AIEd has become a field of scientific research for more than 30 years (Luckin, Holmes, Griffiths, & Forcier, 2016). Interest in understanding and improving the adoption of AI techniques for educational purposes is higher than it has ever been, not only within educational institutions but also in government sectors. There are few reviews of the topic (e.g., Hinojo-Lucena, Aznar-Díaz, Cáceres-Reche, & Romero-Rodríguez, 2019; Roll & Wylie, 2016; Zawacki-Richter, Marín, Bond, & Gouverneur, 2019), yet, to our knowledge, currently, no systematic review of influential research concerning AIEd is available. This research aims to present a thorough overview of the influential academic studies concerning AIEd worldwide using a systematic review method. This study first clarifies what is a broad AIEd and a narrow AIEd, and then conducts a systematic analysis of 45 highly cited AIEd articles to identify the development, trends, and technologies adopted, as well as major research issues concerned by the AIEd community.

This remaining of this section is composed by three subsections. Section 1.1 introduces the application of AI techniques in education. Section 1.2 introduces citation-based literature analysis. Finally, section 1.3 illustrates the research aim and questions.

* Corresponding author.
*E-mail addresses:* xielingchen0708@gmail.com (X. Chen), hrxie2@gmail.com (H. Xie), dizoudaisy@gmail.com (D. Zou), gjhwang.academic@gmail.com (G.-J. Hwang).

https://doi.org/10.1016/j.caeai.2020.100002
Received 18 July 2020; Received in revised form 5 August 2020; Accepted 5 August 2020

### 1.1. Artificial intelligence in education

AI-enhanced digital technology has played an essential role in our daily life, with its great power of changing the way we think, act, and interact. Ever since its emergence, AI has developed with prosperity and flourishment, in particular with the emergence of Artificial Neural Networks (ANN) and Deep Learning (DL) (Chan & Zary, 2019).

As a matter of fact, although the concept of AI and intelligent machines can be dated to the 14th century (Tatar, Roschelle, Vahey, & Penuel, 2003), the idea of AIEd remains much younger for about 25 years, with AI being brought into education area through different channels and in different forms (Heffernan & Heffernan, 2014; Koedinger & Corbett, 2006). The way how we teach and learn has been dramatically influenced by continuingly emerged outcomes and achievements from different disciplines and areas (Humble & Mozelius, 2019), particularly computer science. (Bayne, 2015; Botrel, Holz, & Kübler, 2015).

Applications and tools driven by AI technologies, for instance, intelligent robots and adaptive learning systems, have been increasingly utilized by educators and learners within both K-12 and university contexts. AI technologies provide opportunities for the realization of personalized learning for learners to meet their individual needs (Della Ventura, 2017). Because everyone is independent and has unique learning styles, abilities, and needs, it can be difficult to satisfy every learner by using traditional educational methods. However, with AI, instructors can suit everyone's needs on a case-by-case basis (Della Ventura, 2017). Thus, learners are able to be more motivated, engaged, and independent in the process of learning (Della Ventura, 2018; Wang, 2017). In addition, AI technologies offer chances to support the engagement of learners with learning disabilities.

With the increasingly wide application of AI technologies for teaching and learning, instructors are offered chances to get rid of repetitive and tedious tasks and to reply to students timely, thus advancing the adaptive and personalized teaching process (Chan & Zary, 2019). Particularly, hardware advancements, for instance, the graphical processing units with high-speed and the accessibility to various software libraries, have stimulated the application of AI technologies, in particular with the prosperity of DL research and the implementation of data analysis techniques. Furthermore, essentially, to a large degree, the future development of education will be closely related to the development of AI. Hence, future education will be further stimulated and bloomed with the developments and advancements of novel technologies and computing capacities of intelligent machines.

As displayed in Fig. 1, a total of 143 articles concerning AIEd from six widely-acknowledged SSCI journals (i.e., *Computers & Education*, *Educational Technology Society*, *Interactive Learning Environments*, *British Journal of Educational Technology*, *Educational Technology Research and Development*, and *Journal of Computer Assisted Learning*) related to educational technology have been published during the period 1999–2019. Averagely, for each year, about eight articles in relation to AIEd were published in these journals.

Research on AIEd involves diverse research foci. For instance, Rodrigues and Oliveira (2014) proposed a formative assessment system for learners which was able to create and assess exams and monitor learners' learning progress. By comparing with the manually graded results, it was found the proposed system was effective in grading exams. Eguchi (2016) introduced RoboCupJunior and explored how effective it was for improving the learning of STEM contents and skills among learners. Through the evaluation of learners' subject knowledge, Lan, Waters, Studer, and Baraniuk (2014) proposed an innovative method concerning Machine Learning (ML)-driven learning analytics.

Given the increasing number of AIEd research articles being published, it is necessary and essential to systematically clarify and discuss relevant issues. There are several studies exploring issues regarding the application of AI in education contexts. Some recent relevant studies concerning issues related to AIEd are summarized in Table 1. There are three clear limitations. First, most are position studies in which authors expressed their personal understanding and opinions concerning issues related to AIEd (e.g., Florea & Radu, 2019; Garg, 2020; Ocaña-Fernández, Valenzuela-Fernández, & Garro-Aburto, 2019; Zovko & Gudlin, 2019). Such studies may involve bias and may not objectively reflect the practical situations since they can be subjective without or with little support evidence bibliographically. Second, very limited studies have explored issues concerning AIEd bibliographically. Hinojo-Lucena et al. (2019) presented a bibliometric analysis of 132 academic studies concerning the application of AI in higher education published during the period 2007–2017, to explore the status of production, the major source titles, organizations, authors, and countries. Although bibliometric analysis is principally useful in analyzing sizable literature data, an in-depth investigation by using a systematic review methodology is needed. It is encouraged to review representative articles systematically to understand the domain of AIEd more comprehensively. In addition, although there is a study by Roll and Wylie (2016) which provided a systematic

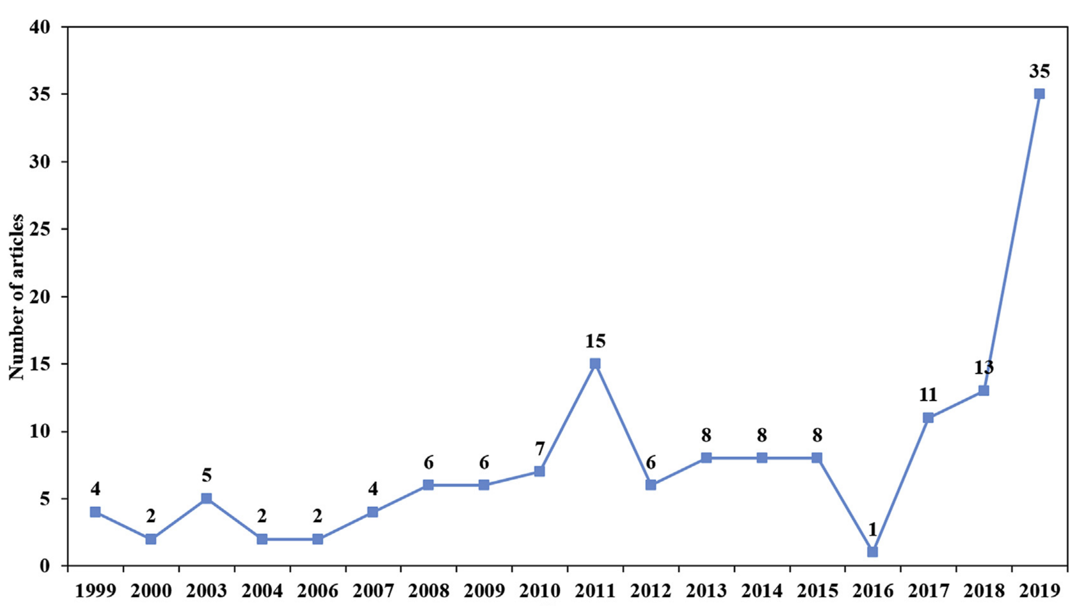



**Fig. 1.** The number of articles on AIEd in six widely-acknowledged SSCI journals related to educational technology during the period 1999–2019.

**Table 1**
Examples of recent studies concerning issues related to AIEd.

| Author(s) and year | Methods | Analysis aspects |
| --- | --- | --- |
| Hinojo-Lucena et al. (2019) | Bibliometric analysis of 132 scientific articles on AI in higher education indexed in Web of Science and Scopus databases during 2007–2017 | To explore the status of production, to investigate the relationship between the number of authors and papers, and to explore the main source titles, organizations, authors, and countries with the highest scientific output on artificial intelligence in higher education. |
| Roll and Wylie (2016) | A systematic overreview of 47 articles from 1994, 2004, and 2014 in the Journal of AIED | To explore the foci and representative scenarios occupied the field of AIED. |
| Garg (2020) | Position paper, providing guidance for medical educators to be properly prepared for AI. | To introduce the broad concepts of AI; to present the impact of AI on medicine; to identify AI's direct impact on the method and content of medical education; to help medical educators get ready for the new demands and chances. |
| Ocaña-Fernández et al. (2019) | Position paper, discussing issues relating to AI and its implications in higher education | To discuss several issues, including AI and its impact on the global world, human intelligence and AI, the traditional university versus the new university, digital skills, intelligent tutoring systems, and online learning, as well as new trends towards a globalized social learning |
| Zovko and Gudlin (2019) | Position paper, discussing issues concerning consequences brought by AI disruptive technology to students, faculty, and society. | To give an overview of disruptive innovations and technologies with the focus on artificial intelligence as a disruptive technology. Special focus is given to the limits and obstacles of the introduction of artificial intelligence in educational processes and the educational system in general. |
| Florea and Radu (2019) | Position paper, discussing issues relating to the relationship between AI and education: | To discuss two views on the relationship between AI and education: one on how can AI enhance education, help personalize the learning experience, help teachers in their endeavor, and its impact on e-learning, and the other on how education in AI has to be conceived to create the workforce required to face this new technological revolution. |

overreview of academic literature, it only considered articles published by the *International Journal of Artificial Intelligence in Education* in 1994, 2004, and 2014. There are two problems regarding the study. First, with only a single journal, findings obtained may not provide an essential understanding of the research area of AIEd. Second, studies that have been published after 2014 have yet to be reviewed. Further, with the recent rapid technological advances and innovations, such review already appears dated. As a result, there have been few, if any, in-depth reviews summarising the trends, developments, and innovative AI technologies used for educational purposes. Zawacki-Richter et al. (2019) synthesized 146 articles regarding the application of AI in higher education and described four broad areas of AIEd applications, including profiling and prediction, intelligent tutoring systems (ITSs), assessment and evaluation, as well as adaptive learning systems. However, the study did not select articles for analysis from an academic influence and impact perspective.

To fill in these gaps, the present study reviews the scientific articles related to AIEd, particularly those influential ones indexed by SCI and SSCI, and are highly cited over the period from 1999 to 2019 to explore important issues that remained to be investigated. For instance, what was the annual distribution of impactful AIEd articles? Which important actors had contributed most to impactful AIEd articles? What were the focuses and hotspots concerned by impactful AIEd articles? What were the major AI technologies adopted by impactful AIEd research? To solve these questions, a systematic analysis of highly cited AIEd articles is suitable.

### 1.2. Analysis of highly cited research articles

Citation-based literature analysis provides analyses of scientific literature from qualitative and quantitative perspectives to identify articles with the most influence on a specific research area (Garfield & Merton, 1979; Godin, 2006). As indicated by Brace (1992), the citations that an article received could effectively reflect its quality. Levitt and Thelwall (2008) also highlighted the relationship between the highly cited studies with the research of high quality. Highly cited articles are capable of revealing research development to present the progress of science (Baltussen & Kindler, 2004; Ohba, Nakao, Isashiki, & Ohba, 2007). In this sense, highly cited articles can, therefore, be adopted to analyze historical trends and emerging research themes, as well as to serve as guidance for future research (Choudhri, Siddiqui, Khan, & Cohen, 2015). In addition, highly cited articles provide a measure of the productivity of researchers and institutions from an objective perspective, helping identify predominant actors contributing to a particular research area (Abramo & D'Angelo, 2011).

Citation-based literature analysis has been popularly used to evaluate research fields of different disciplines. For instance, Liao, Tang, Li, and Lev (2019) carried out an analysis of 646 Essential Science Indicators, highly cited articles relating to operations research and management science published from 2008 to 2017. They conducted analyses in different aspects, including the most influential actors, the scientific collaboration relationship, research topics, and future directions. To examine the research status and trend of research concerning immunotherapy for childhood leukemia, Zhong et al. (2019) carried out an analysis of 100 most-cited articles to provide implications about research foci. Using both bibliography and survey data, Confraria, Blanckenberg, and Swart (2018) identified researchers that had contributed the most. They further conducted an econometric analysis to examine what characteristics contributed to being highly cited. Results indicated that averagely, authors contributing more scientific articles annually, cooperating more with non-African authors, and receiving the highest qualification from an Anglo-Saxon academic institution, tended more to be highly cited.

In particular, the citation-based literature analysis is also widely utilized in educational disciplines. For instance, based on Web of Science (WoS) and resume or CV data, Martinez & Sá (2019) explored the career trajectories and publication patterns of the most-cited authors affiliated with Brazilian academic institutions. Findings obtained highlighted the international mobility of the authors from the early phases of academic careers. Lai (2019) reviewed the 100 most impactful mobile learning studies to examine the latest trends in mobile learning research. Major research directions, for example, proposing novel types of learning strategies, and utilizing mobile learning to wider disciplines and domains, were uncovered. Arden, Pentimonti, Cooray, and Jackson (2018) utilized categorical content analysis to examine major research trends and issues covered within highly cited articles from special education journals. Results obtained demonstrated the research trends in articles of commentary and position types, reading and behavior, as well as experimental and quasi-experimental designs with small proportions.

### 1.3. Research aims

Motivated by the flourishing development of AIEd research, it is worthy of clarifying its definition as well as its relationship with other similar areas, including Educational Data Mining (EDM), Computer-Based Education (CBE), and Learning Analytics (LA). In addition, driven by the increasing number of AIEd articles, it can be essential to carry out a systematic review of AIEd to uncover the most influential research studies as well as their distribution, development, major contributors, and research foci. However, in accordance with our literature investigation, no systematic analysis has concentrated specifically on influential AIEd research literature, a research area worth exploring. This paper is built on two previous reviews on AIEd research in which scientific production on AI in higher education was analyzed using bibliometric analysis (Hinojo-Lucena et al., 2019) and research foci that occupied the Journal of AIED have been investigated using systematic analysis (Roll & Wylie, 2016). The goal of the current systematic analysis, as the first in-depth work, is to systematically and comprehensively review influential research studies on AIEd published by SSCI- and SCI-indexed journals to understand the overall applications of AI technologies in educational contexts. To realize the goal, this study first explored the annual distribution of influential AIED studies and then identified the major contributors in terms of leading journals, institutions, and countries/regions. In this way, we facilitate the understanding of "what had the status of production been over time" and "what were organizations and countries with the highest scientific output on AIEd" (Hinojo-Lucena et al., 2019). Second, we facilitate the understanding of "what were the current foci of research in AIED" (Roll & Wylie, 2016) by exploring the most frequently used terms in the highly cited AIEd studies. Third, by systematically examining analyzing the AI technologies adopted in each of the AIEd studies, this review study can assist scholars in understanding what types of and to what extent AI techniques have been adopted within the current AIEd research work that has received wide attention and brought significant influence, as well as suggestions about future directions such as the potential novel technologies and applications of AIEd (Chen, Zou, Cheng, & Xie, 2020). Furthermore, we provided the definition of AIEd from broad and narrow perspectives by clarifying the relationship between AI, ML, and DL, so as to provide a criterion for better classifying AIEd studies, as well as monitoring the research development stage. In addition, we clarified the relationship between AIEd, EDM, CBE, and LA to help understand what subject attributes each concept relates to, as well as recognizing overlaps between them. In short, this review is indispensable and essential to link existing AIEd research and future trend of the application of AI technologies for educational purposes.

Specifically, there are five research questions to be investigated:

RQ1: How did broad and narrow AIEd studies distribute?

RQ2: What journals, institutions, and countries/regions contributed the most to the highly cited AIEd studies?

RQ3: What were the major research issues and AI technologies adopted in the highly cited AIEd studies?

RQ4: What theories and frameworks had been used in the highly cited AIEd studies?

The significance of our study to the research field of AIEd can be summarized as follows. First, it provides scholars devoting to the area of educational technology to understand the status and trend of AIEd research, particularly in terms of influential studies. Second, it assists researchers to recognize the most impactful AIEd studies and be more aware of what types of AI technologies and what types of contexts can them be adopted to assist the achievement of educational purposes. Third, to appreciate influential institutions and countries/regions that are potential AIEd research collaborators. Third, it raises researchers' awareness of the research foci, impactful research issue. Fourth, it helps newcomers to the field in terms of recognizing theories and technologies that are commonly adopted in high cited AIEd studies. Fifth, scholars can recognize influential journals in publishing AIEd studies. In addition, the findings obtained are also useful in terms of resource allocation carried out by decision-makers, for example, educational institutions and governments, in designing and conducting scientific, technological, and educational activities.

## 2. Definition of Artificial Intelligence in Education

### 2.1. Artificial intelligence, machine learning, deep learning

The research field of Artificial Intelligence (AI) began to serve as an academic discipline since 1956 (Knapp, 2006), and thereafter, AI has undergone several waves of optimism, being partitioned into various sub-fields. Such sub-fields are divided primarily according to technical considerations like specific goals, for example, Machine Learning (ML), and specific tools, for example, "logic" or "ANNs." Fig. 2 depicts the major evolutionary process of AI, from initial AI, to ML, and to the recent Deep Learning (DL).

#### 2.1.1. Artificial intelligence

It can be difficult to give an exact definition to Artificial Intelligence (AI), even for domain experts, owing to two major reasons. For one thing, what AI involves is continuously evolving. As indicated by Nick Bostrom, various leading-edge AI technologies have been incorporated in universal applications and are not named AI (Bostrom, 2006). For another thing, AI is itself an interdisciplinary area with researchers and experts from a variety of areas, for example, neuroscience, psychology, and linguistics, constantly making contributions by bringing their own perception, knowledge, and terminology.

There have been many scholars attempting to define AI. For instance, Russell and Norvig (2016) indicated that AI was used for describing machines or computers imitating "cognitive" functions, for example, "learning" and "problem solving," associated with the human mind. Poole, Goebel, and Mackworth (1998) considered AI the study of intelligent agents that were able to perceive their settings as well as achieving a particular goal by maximizing the probability. Kaplan and Haenlein (2019) regarded AI as the ability of a system to precisely interpret and learn from inputted data, and to further utilize what it learns to achieve a particular goal. Although there is not a commonly acknowledged definition of AI, what can be sure is that besides computer science, AI also involves disciplines such as information science, psychology, linguistics, neuroscience, philosophy, mathematics, and many others. AI is thus an interdisciplinary area with strong comprehensiveness (H. Zhao, Li, & Feng, 2018). In this study, we followed the definition of AI by Baker and Smith (2019) as "computers which perform cognitive tasks, usually associated with human minds, particularly learning and problem-solving (p. 10)" (Baker and Smith, 2019).

#### 2.1.2. Machine learning

Evolved from AI, particularly from the research regarding pattern recognition and computational learning (Karmani, Chandio, Korejo, & Chandio, 2018), Machine Learning (ML) has a primary aim as to propose methods enabling learning from data and to further make predictions (Kohavi & Provost, 1998). ML was initially defined by Arthur Samuel (1959) as the study of how to make learning without being specifically programmed. In addition, a broadly cited, more formal definition of ML was offered by Tom M. Mitchell (1997), where ML serves as "a computer program is said to learn from experience E with respect to some class of tasks T and performance measure P if its performance at tasks in T, as measured by P, improves with experience E (p. 2)."

ML's power lies in its ability to surmount severely static program instructions by modeling on the inputted data and making data-enhanced predictions or decisions. It is universally acknowledged that ML is a technique for training intelligent agents to accomplish a particular task using large volumes of data (e.g., Brynjolfsson & Mitchell, 2017; Gori, 2017; Murphy, 2012). Similarly, AI aims to build intelligent applications of machine level, a core mission of which is to propose behavior systems that are able to imitate the human brain and are controlled by a human-computer system (Han, 2018).

The ML algorithms mainly involve clustering, classification, decision tree learning, reinforcement learning (RL), inductive logic programming, and Bayesian networks. According to Jordan and Mitchell (2015), within

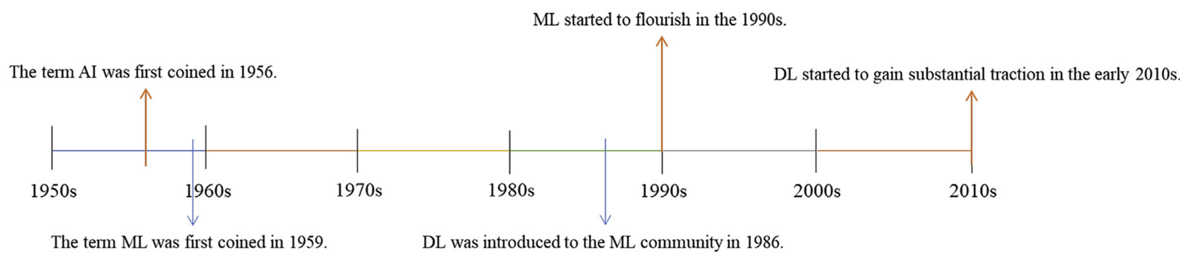



**Fig. 2.** The major evolutionary process of AI (Note: developed based on the introduction of AI in Wikipedia[11])

AI, ML has emerged to be a choice for developing practical tools to deal with various issues, for example, speech recognition and natural language processing (NLP). Nevertheless, to our knowledge, none of the ML algorithms has accomplished the ultimate goal of general AI. In addition, with only the early ML algorithms, even narrow AI is mostly out of reach.

As a subset of AI, ML is also of interdisciplinary nature. In line with Michael I. Jordan, the concept of ML has rooted in statistics from methodological principles to theoretical applications. In addition, there are also a variety of other disciplines involved. According to Jordan and Mitchell (2015), ML "sits at the crossroads of computer science, statistics and a variety of other disciplines concerned with automatic improvement over time, and inference and decision-making under uncertainty (p. 2)."

*2.1.3. Deep learning*

The term Deep Learning (DL) was first introduced by Rina Dechter to the ML community in 1986 (Dechter, 1986; Schmidhuber, 2015), and to the ANNs by Igor Aizenberg and his colleagues in 2000 (Aizenberg, Aizenberg, & Vandewalle, 2013; Gomez & Schmidhuber, 2005). As part of ML algorithms (Deng & Yu, 2014), DL focuses on extracting higher-level features from the inputted data by adopting multiple layers. According to LeCun, Bengio, and Hinton (2015), as one of the most efficient ML techniques, DL algorithms can be categorized as supervised, semi-supervised, and unsupervised (Bengio, Courville, & Vincent, 2013; LeCun, Bengio, & Hinton, 2015; Schmidhuber, 2015).

As indicated by the moniker "neural networks (NNs)," ANNs is an inspiration of biological structures. However, neural networks differ from biological brains in many aspects. For example, different from the biological brain where most of the living organisms are dynamic and analog, NNs are more static and symbolic (Bengio, Lee, Bornschein, Mesnard, & Lin, 2015; Marblestone, Wayne, & Kording, 2016; Olshausen & Field, 1996). Thanks to the advancements of neural networks over the past few years, AI has become more powerful and effective (LeCun et al., 2015; Schmidhuber, 2015). In addition, the successful applications of neural networks for AI owe much to neuroscience, which has offered initial guidance toward both architectural and algorithmic constraints (Hassabis, Kumaran, Summerfield, & Botvinick, 2017).

The relationship between DL, ML, and AI can be depicted as Fig. 3, which shows that DL serves as ML's subset, and ML serves as AI's subset. In this sense, we define AIEd from both broad and narrow perspectives. The broad AIEd is the adoption of AI techniques in education. There are two narrow definitions for AIEd. The first one is MLEd, which is the adoption of ML techniques in education, while the second one is DLEd, referring to the adoption of DL techniques education.

*2.2. Artificial intelligence in education, educational data mining, computer-based education, learning analytics*

Artificial Intelligence in Education (AIEd) can be considered to be the combination of three major fields, including computer science, statistics, as well as education. In addition to these three areas, AIEd is also an interdisciplinary field involving but not limited to, for example, cognitive psychology and neuroscience. The intersection of the three major fields also generates other sub-fields that are in close relation to AIEd, for example, Educational Data Mining (EDM), Learning Analytics (LA), and Computer-Based Education (CBE) (See Fig. 4).

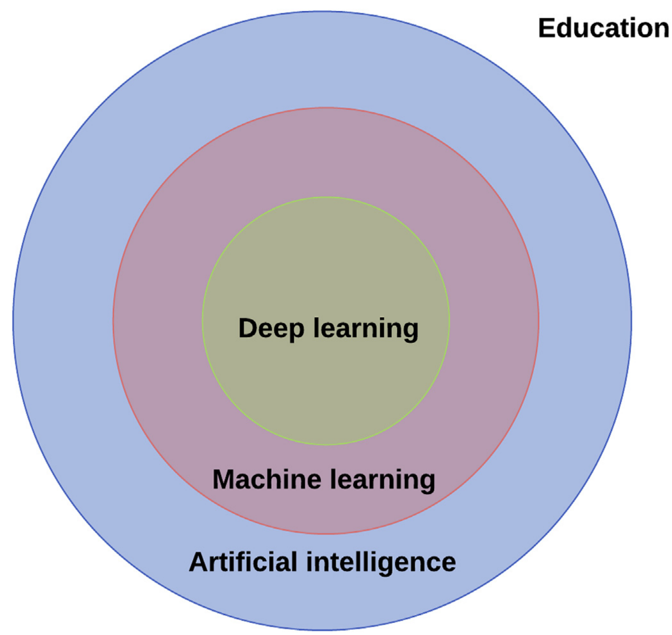



**Fig. 3.** Concepts of AIEd from broad and narrow perspectives adapted from (Sze, Chen, Yang, & Emer, 2017).

*2.2.1. Educational data mining*

Educational Data Mining (EDM) is the analysis of various types of educational data by using statistical, ML, and DM algorithms (Romero & Ventura, 2010). It focuses on the creation of approaches to examine the unique educational data, so as to understand the way learners learn and identify the settings where they obtain higher learning outcomes as well as deeper insights into and understanding of educational phenomena (Baepler & Murdoch, 2010). Studies regarding EDM include two different types, namely, statistics and visualization, as well as web mining involving technologies like clustering, classification, association rule mining, and text mining (Romero & Ventura, 2007).

As indicated by Baker and Yacef (2009), new types of EDM studies would emerge to be related to the exploration with models, an approach which has been prominent in cognitive modeling and bioinformatics research but is rare in the research of education. In this sense, EDM approaches can induce some level of influence on education and its relevant interdisciplinary areas such as AIEd. A multitude of applications driven by AIEd has already been adopted in many universities and other educational institutions, many of which integrate AIEd and EDM techniques to "track" students' behaviors, for instance, identifying students' risk at abandoning their studies so as to provide timely support through the analysis of data regarding class attendance and assignment submission (Luckin et al., 2016).

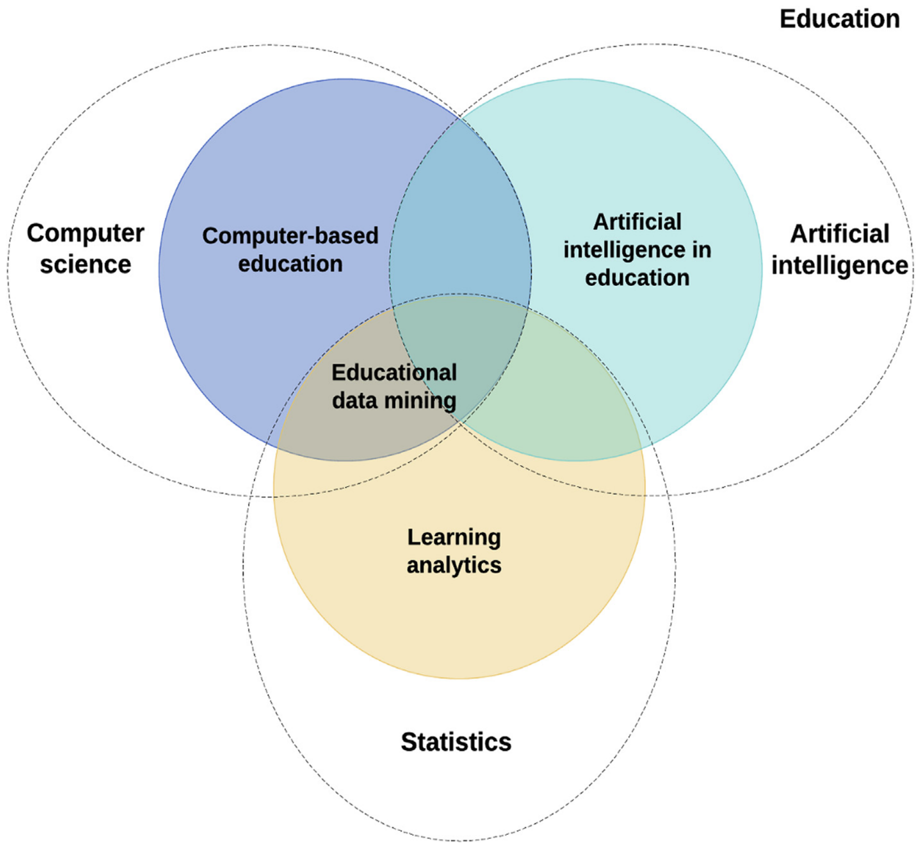



**Fig. 4.** The relationship between AIEd, EDM, CBE, and LA adapted from (Romero and Ventura, 2013).

### 2.2.2. *Learning analytics*

Learning Analytics (Baepler & Murdoch, 2010) is considered as "the measurement, collection, analysis, and reporting of data about learners and their contexts, for purposes of understanding and optimizing learning and the environments in which it occurs (Romero & Ventura, 2013) (p. 2)." There are similarities in attributes, interests, and goals for Learning Analytics (LA) and EDM. However, they have major differences mainly regarding techniques and emphasis as follows (Siemens & Baker, 2012). For one thing, the most adopted techniques in LA involve statistics, visualization, discourse analysis, social network analysis, and sense-making models. However, the most adopted algorithms in EDM involve clustering, classification, Bayesian modeling, relationship mining, as well as discovery with models. For another thing, LA concerns more about describing data and results, while EDM emphasizes much on describing and comparing the DM technologies.

### 2.2.3. *Computer-based education*

Computer-Based Education (CBE) is defined as the use of computers in education for providing instructions to learners. Initially, CBE systems were individual tools running on a local computer with no use of AI to address issues such as student modeling, adaptation, and personalization. With the universal adoption of the Internet, new educational web-driven systems like e-learning systems have emerged. In addition, the increasing use of AI techniques has stimulated novel types of adaptive and intelligent systems for educational purposes. Hence, there are some overlaps between CBE and AIEd, with several major types of CBE systems adopted currently are also AIEd systems, for instance, ITS (Mostow & Beck, 2006), learning management system (Romero, Ventura, & García, 2008), adaptive hypermedia and multimedia system (Merceron & Yacef, 2004), test and quiz system (Romero, Zafra, Luna, & Ventura, 2013), and ubiquitous learning environment.

## 3. A systematic review of the highly cited AIEd research

The statement on the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (Moher, Liberati, Tetzlaff, & Altman, 2009) was considered when conducting this review and reporting on it. Fig. 5 depicts the framework of article collection.

### 3.1. *Search strategy*

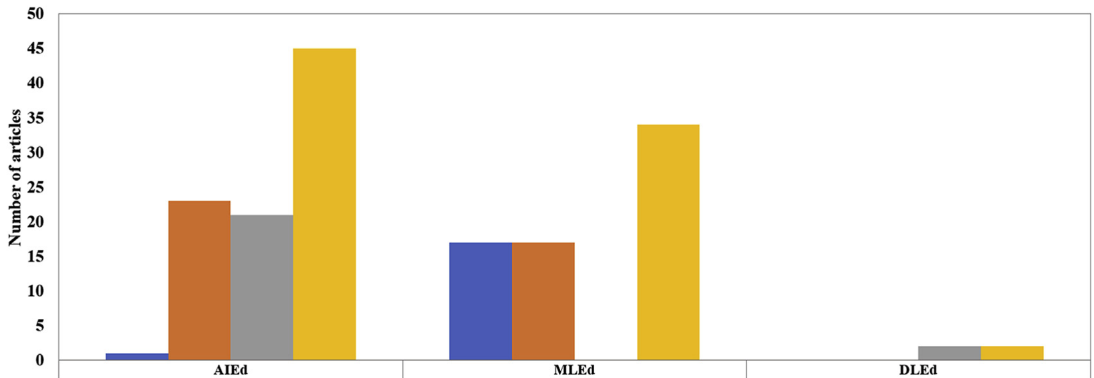
In conducting data retrieval, we used keywords list of AI developed by existing studies (Tran et al., 2019), for example, "artificial intelligence," "artificial neural network*," "machine intelligence," "machine learning," "deep learn*," and "robotic*," as search terms. As for keywords list of education, by referring to several bibliometric studies (e.g., Hallinger & Chatpinyakoop, 2019; Heradio et al., 2016; Martí-Parreño, Méndez-Ibáñez, & Alonso-Arroyo, 2016; Sweileh, Al-Jabi, Zyoud, & Sawalha, 2018) in relation to education, we finalized to include "education" and "teaching," while "learning" was excluded for the reason that it might cause ambiguity with some articles containing "learning" but are irrelevant to education, for example, articles related to ML in other areas such as medicine.

With the Internet being so influential in research over the last 20-year period (Natividad, Spector, & Evangelopoulos, 2018), the field of educational technology has experienced rapid growth. Thus, we examined two decades of publications, namely, the years 1999–2019.

As indicated by several review studies (e.g., Chen, Zou, & Xie, 2020; Chen et al., 2018), it is essential to choose studies indexed by high-quality databases. Thus, the WoS database[2] was selected for this

[1] https://en.wikipedia.org/wiki/Artificial_intelligence.

[2] https://webofknowledge.com.




| | AIEd | MLEd | DLEd |
|---|---|---|---|
| 1990-2000 | 1 | 17 | 0 |
| 2001-2009 | 23 | 17 | 0 |
| 2010-2016 | 21 | 0 | 2 |
| Total articles | 45 | 34 | 2 |

**Fig. 5.** A framework of article selection.

study to collect raw bibliographic data in relation to the research area of AIEd, since WoS a reputable journal article collection (Cobo, Martínez, Gutiérrez-Salcedo, Fujita, & Herrera-Viedma, 2015), We used "the advanced query as follows, where TS" (Topics) was a search field. The search was conducted on 1st January 2020.

TS=(("Artificial intelligence" OR "Machine intelligence" OR "artificial neural network*" OR "Machine learning" OR "Deep learn*" OR "Natural language process*" OR "Robotic*" OR "thinking computer system" OR "fuzzy expert system*" OR "evolutionary computation" OR "hybrid intelligent system*") AND ("education*" OR "teaching"))

### 3.2. Study selection

Articles collected were included if they matched all the following criteria. First, it should be of Science Citation Index or Social Sciences Citation Index. Second, it is written in English. Third, it was published from 1970 to 2019. Finally, it must be a research article with original contributions (Geng et al., 2017). Following the criteria, we finalized and downloaded 1,535 articles with their bibliographic data and annual citation data. Elements for each publication, for example, title, publication source, year of publication, abstract, and author address, were summarized and organized in Excel for data processing.

Many studies (e.g., Ho & Hartley, 2017; Ivanović & Ho, 2019; Zhong et al., 2019) used highly cited articles to conduct analysis, by sorting articles in descending order according to citation count, and then selected the influential ones. As proposed by Chuang, Wang, and Ho (2011) and Ming, Hui, and Yuh (2011), $TC_{year}$ is defined as the total citations received by an article since its publication to a selected year. In this study, to guarantee repeatability (Fu, Wang, & Ho, 2012), $TC_{2019}$ was adopted. $TC_{2019}$ is defined as the total citations an article has received since its publication to the end of 2019. $TC_{2019} >= 20$ was employed to select 257 articles to be highly cited articles. We then conducted filtering to exclude those irrelevant to both the narrow AIEd and the broad AIEd, and for those relevant ones, we conducted coding to determine whether it is DLEd, MLEd, or AIEd.

To assure the closeness of the collected articles to AIEd and enhance analysis efficiency, a filtering work was carried out by two domain experts. They carefully read and analyzed each article in accordance with the listed criteria in Table 2 to narrow down the collected articles. Each article must be relevant to AI-enhanced education. Specifically, each article should relate to the use of DL, ML, or other AI techniques to back up teaching or learning. If not, then it would be excluded. For instance (Laird, Shoup, Kuh, & Schwarz, 2008), explored how discipline affected learners' adoption of and instructors' emphasis on deep learning, as well as the relationship between educational outcomes and deep learning. However, in their study, deep learning donated learners' engagement in approaches to learning with an emphasis on the integration, synthesis, and reflection, rather than representing the AI techniques. Thus, it was not included in this study.

Following the criteria, two domain experts coded the 257 highly cited AIEd articles, respectively. We adopted Cohen's kappa coefficient $\kappa$ (McHugh, 2012) as a measure of inter-rater reliability to see how much consensus existed in the judgment given by the two experts. Given the two experts each categorized $N$ items into $C$ categories, $\kappa$ is defined as $\kappa = \frac{p_o - p_e}{1 - p_e}$, in which $p_o$ donates the observed agreement between the two experts, and $p_e$ donates the hypothetical possibility of agreement. If there are $k$ categories, $N$ items with $n_{ki}$ representing the number of times expert $i$ predicts category $k$, then $p_e$ is calculated as $p_e = \frac{1}{N^2}\sum_k n_{k1} n_{k2}$. Through calculation, the Cohen's kappa coefficient $\kappa$ of the two domain experts was 94%. As for the inconsistently coded articles, the experts conducted a discussion to reach an agreement. Using the criterion, 207 articles were excluded, including 157 publications irrelevant to AI, and 49 publications irrelevant to education, and six are of review-typed, leaving 45 highly cited articles relevant to AIEd.

**Table 2**
Inclusion and exclusion criteria.

| | | |
|---|---|---|
| Inclusion criteria | I1 | AI technique use for assisting teaching |
| | I2 | AI technique use for assisting learning |
| | I3 | AI technique use for e-learning systems |
| | I4 | AI technique use for pedagogical innovations |
| | I5 | AI technique use for educational data mining |
| | I6 | AI technique use for learning analytics |
| | I7 | AI technique use for domain-specific applications in education |
| | I8 | AI technique use for educational data analysis |
| | I9 | AI technique use for student/learner modeling |
| Exclusion criteria | E1 | AI courses learning |
| | E2 | No mention of AI techniques uses in education |
| | E3 | Not related to education |
| | E4 | The mention of deep learning is not about AI techniques, but about learning deeply |
| | E5 | Subjects for STEM education (virtual reality, robotics, programming) |
| | E6 | AI or ML related courses or teaching methods |
| | E7 | Computational thinking related studies |
| | E8 | E-learning systems without the use of AI techniques |
| | E9 | Computer science courses learning |

### 3.3. Data extraction and quality synthesis

Data were extracted from each article on (a) the study characteristics, including author, authors' institutions and countries/regions, publication sources, year of publication, citation count and author defined keywords, and (b) application scenario type, research issues, adopted AI technologies, adopted theories, proposed theoretical or methodological frameworks, as well as definitions concerned (DLEd, MLEd, or AIEd). In addition, important terms were extracted from titles and abstracts with the use of the Natural Language Toolkit. Two domain experts piloted coding for the selected 45 highly cited articles according to their full-text content to determine whether they were in relation to DLEd, MLEd, or AIEd, by examining the adopted techniques to conduct the categorization. Some examples of the popular adopted techniques of DL, ML, and AI are shown in Table 3. To be specific, the coding was conducted following three steps. First, if the article was in relation to the use of DL techniques, for example, ANNs, then it was coded as DLEd. Second, if the article was not in relation to the use of DL techniques, but was in relation to the use of ML techniques, for example, Bayesian networks, then it was coded as MLEd. Third, if there was no mention of DL or ML techniques, then it would be coded as AIEd. For example, Thomaz and Breazeal (2008) introduced a simulated RL robot and analyzed real-time human teaching behavior. This article was related to "teaching an RL agent." Thus, it was coded as MLEd. Following the three-step strategy, the two domain experts coded all the 45 articles, respectively, with inter-rater reliability above 95%.

## 4. Results

Table 4 summarizes the major features of the 45 influential AIEd studies in terms of journal, citation count, authors' regions and institutions, application scenario type, research issues, adopted AI technologies, as well as definitions concerned (DLEd, MLEd, or AIEd). Table 5 summaries the education-related theories that were clearly specified in the studies, as well as novel models or frameworks proposed in the studies. Detailed descriptions and discussions are described in the results and discussion sections.

### 4.1. Distributions of AIEd, MLEd, and DLEd

The distributions of AIEd, MLEd, and DLEd in terms of article count over three periods (i.e., 1990–2000, 2001–2009, and 2010–2016) is shown in Fig. 6. Among the 45 highly cited AIEd articles, there are 41 articles belonging to MLEd, while only two belonging to DLEd. The first study is by Kasabov (2014), which proposed an analytical model for the prediction and explanation of freshmen student attrition by using data mining technologies. Their experimental results indicated that the educational and financial variables were essential reasons for student attrition. The second study is by Tan et al. (2012), which introduced an intelligent computer-supported instruction modeling approach for instructing parallel robot. With reference to the educational profile, they analyzed the cognitive processes of mechatronics students when they were learning parallel robots. They also conducted comparisons about the advantages of different approaches in teaching the same topic. Two terms of teaching experiments using the proposed approach have demonstrated its effectiveness in enhancing students' learning efficiency as well as their comprehension and mastery of the working principles of a parallel robot.

The figure shows that in the first period, namely 1990–2000, there were very limited articles found to be highly cited. Most of the influential AIEd articles were published during the latter two periods, namely, 2001–2009, and 2010–2016. The results indicate a flourishing development of AIEd since the 2000s, as well as a high impact and influence of the articles published during that period.

**Table 3**
Examples of the popular adopted techniques of DL, ML, and AI.

| Techniques | Examples of algorithms |
|---|---|
| DL | artificial neural networks, deep belief networks, deep neural networks, recurrent neural networks, convolutional neural networks, long short-term memory, generative adversarial networks, variational auto-encoders, adversarial learning, convolutional neural networks |
| ML | reinforcement learning, decision trees, support vector machines, regression analysis, Bayesian networks, genetic algorithms |
| AI | fuzzy logic, rule-based systems, agent system, heuristic algorithms |

### 4.2. Article count and citation count of AIEd

The distributions of article count, citation count, and citations per article of the highly cited AIEd research are shown in Fig. 7. None of 45 highly cited AIEd articles were published before the year 1990. No research article published after the year 2016 has been found to be highly cited. The year with the most articles was 2010. The first two research articles found to be highly cited were published in 1996 by Miyamoto et al. (1996). Furthermore, the time period 2008–2010 concentrated most of them. It is noteworthy that the ranking position of the most recently published highly cited AIEd article was contributed by Eguchi (2016). Moreover, although highly cited articles are typically located in remotest years because of the citation window, AIEd is currently growing fast, as research articles are more likely to catch the attention of scholars within years of being published. From the results, we can see that there is an obvious growth in citations, indicating an increasing influence of the highly cited AIEd articles.

### 4.3. Top journals

The 45 influential AIEd articles were contributed by a total of 30 journals, among which the top ten are depicted in Table 6. *Computers & Education* published seven of these articles, followed by *Expert Systems with Applications* with three articles, and their impact factors in the year 2018 were 5.902 and 4.577, respectively. As expected, most of the influential AIEd articles were distributed in journals with high impact factors.

Fig. 8 shows the article distributions of the top ten journals for publishing the 45 highly cited articles over three periods. For all the ten journals, most of their influential AIEd articles were published during the later periods, indicating a recent flourishing development of the AIEd area. In particular, *Computers & Education* had four articles published during the period 2001–2009. For journals such as *Expert Systems with Applications*, all its highly cited AIEd articles were published during the last period.

### 4.4. Top institutions, and countries/regions

A total of 19 countries/regions and 68 institutions have contributed to the publication of the 45 highly cited AIEd articles, the top nine of which are listed in the left part of Table 7. The USA contributed 18 articles, followed by Greece with eight, as well as Spain with four. It is worth noting that although with only two articles, the articles contributed by Argentina had great influence and impact with an index of citation per article as 104. The right part of Table 7 depicts the top nine institutions in publishing the 45 highly cited AIEd articles. *Stanford University* and *University of Patras* had the most highly cited articles, each with three.

Figs. 9 and 10 show the article distributions for the highly cited countries/regions and institutions over three periods. From countries/regions perspective, most of their influential AIEd articles were published in later periods. Particularly, the USA had eight articles published during the period 2001–2009 and nine during the period 2010–2016. For countries such as Spain, all its highly cited AIEd articles were found to be published during the last period, while for countries/regions such as Argentina, Hong Kong, and Taiwan, all their highly cited AIEd articles were found to be published during the second period.

**Table 4**
Summary of the major features of the 45 influential AIEd studies.

| Author(s) and year | Journal | C | Institution(s) | Region(s) | Type | Aspect | AI techniques | DLEd | MLEd | AIEd |
|---|---|---|---|---|---|---|---|---|---|---|
| Chen, Lambert, and Guidry (2010) | Computers & Education | 193 | Univ N Texas; Indiana Univ Bloomington | USA | Online | Student engagement and self-reported learning outcomes | Hierarchical linear model | | ✓ | ✓ |
| García, Amandi, Schiaffino, and Campo (2007) | Computers & Education | 188 | Univ Natl Ctr Provincia Buenos Aires | Argentina | Web | Learning styles | Bayesian networks | | ✓ | ✓ |
| Biswas, Leelawong, Schwartz, & Vye (2005) | Applied Artificial Intelligence | 160 | Vanderbilt Univ; Stanford Univ | USA | Web | Teachable agent | Teachable agent | | | ✓ |
| Thomaz and Breazeal (2008) | Artificial Intelligence | 119 | Mass Inst Technol; Georgia Inst Technol | USA | Web | Teachable robots | RL | | ✓ | ✓ |
| Pierre-Yves (2003) | International Journal of Human-computer Studies | 116 | Sony | France | NS | Personal robots | ML (neural networks, SVM, decision trees) | | ✓ | ✓ |
| Lykourentzou, Giannoukos, Nikolopoulos, Mpardis, and Loumos (2009) | Computers & Education | 97 | Natl Technol Univ Athens | Greece | Online | Dropout prediction | Feed-forward neural networks, SVM | | ✓ | ✓ |
| Miyamoto et al. (1996) | Neural Networks | 83 | Toyota Motor Co Ltd et al. | Japan; USA | NS | Learning robot | Neural networks | | | ✓ |
| Whitehill, Serpell, Lin, Foster, & Movellan (2014) | IEEE Transactions on Affective Computing | 82 | Virginia State Univ et al. | USA | Offline | Student engagement | Binary classification | | ✓ | ✓ |
| Burstein, Chodorow, and Leacock (2004) | AI Magazine | 63 | Educ Testing Serv et al. | USA | Online | Essay evaluation | NLP | | ✓ | ✓ |
| Kotsiantis, Patriarcheas, & Xenos (2010) | Knowledge-Based Systems | 62 | Hellenic Open Univ | Greece | Web | Predict a student's performance | Naive Bayes, nearest neighbor, linear online algorithm | | ✓ | ✓ |
| Calvo, O'Rourke, Jones, Yacef, and Reimann (2010) | IEEE Transactions on Learning Technologies | 57 | Univ Sydney | Australia | Online | Collaborative writing | ML; NLP | | ✓ | ✓ |
| Delen (2010) | Decision Support Systems | 55 | Oklahoma State Univ | USA | NS | Student retention management | ANN, decision trees, logistic regression, SVM, ensembles/ bagging (random forest) | ✓ | ✓ | ✓ |
| Beligiannis, Moschopoulos, Kaperonis, & Likothanassis (2008) | Computers & Operations Research | 48 | Univ Patras, Univ Ioannina | Greece | NS | School timetabling | Evolutionary algorithm | | | ✓ |
| Kotsiantis, Pierrakeas, & Pintelas (2003) | Knowledge-Based Intellignet Information And Engineering Systems | 48 | Univ Patras | Greece | Web | Student dropout prediction | Naive Bayes | | ✓ | ✓ |
| Bittencourt, Costa, Silva, & Soares (2009) | Knowledge-Based Systems | 47 | Univ Fed Campina Grande; Fed Univ Alagoas | Brazil | Web | Web-based educational systems | Semantic web | | | ✓ |
| Kim et al. (2009) | IEEE Computational Intelligence Magazine | 38 | Korea Adv Inst Sci & Technol | South Korea | Web | Robot soccer system | Evolutionary algorithm | | | ✓ |
| Wang et al. (2008) | Computers & Education | 37 | Natl Taiwan Normal Univ et al. | Taiwan; USA | NS | Assessment of creative problem-solving | Regression; least squares; SVM classification | | ✓ | ✓ |
| Lan et al. (2014) | Journal of Machine Learning Research | 36 | Rice Univ; Cornell Univ | USA | NS | Assignments grading | Sparse factor analysis; regression; Bayesian latent factor analysis | | ✓ | ✓ |
| Xu and Wang (2006) | Decision Support Systems | 36 | Univ Queensland; City Univ Hong Kong | Australia; Hong Kong | Web | Autonomous intelligent agents | Fuzzy logic | | ✓ | ✓ |
| Cavus (2010) | Advances in Engineering Software | 34 | Near East Univ | Turkey | Web | Learning management systems | Fuzzy logic | | | ✓ |
| Zouaq & Nkambou (2008) | IEEE Transactions on Learning Technologies | 34 | Univ Quebec Montreal | Canada | Web | Production of personalized e-learning resources | Knowledge extraction; syntactic analysis | | ✓ | ✓ |
| Cheung, Hui, Zhang, and Yiu (2003) | Journal of Systems and Software | 33 | Univ Hong Kong; Stanford Univ | Hong Kong; USA | Web | Intelligent tutoring system | Expert model | | ✓ | ✓ |

(*continued on next page*)

**Table 4** (*continued*)

| Author(s) and year | Journal | C | Institution(s) | Region(s) | Type | Aspect | AI techniques | DLEd | MLEd | AIEd |
|---|---|---|---|---|---|---|---|---|---|---|
| Chen et al. (2010) | Etr&d-educational Technology Research and Development | 31 | Stanford Univ, SRI Int | USA | Web | Teachable agents | Teachable agent | | | ✓ |
| Moridis and Economides (2009) | Computers & Education | 31 | Univ Macedonia | Greece | Online | Prediction of student's mood | Neural networks | | ✓ | ✓ |
| Muldner, Burleson, Van de Sande, and VanLehn (2011) | User Modeling and User-adapted Interaction | 30 | Arizona State Univ | USA | Web | Gaming behaviors analysis | Bayesian network | | ✓ | ✓ |
| Kotsiantis (2012) | Artificial Intelligence Review | 29 | Univ Patras | Greece | Web | Forecasting students' grades | Regression | | ✓ | ✓ |
| Hämäläinen and Vinni (2006) | Intelligent Tutoring Systems, Proceedings | 29 | Univ Joensuu | Finland | Online | Course success prediction | ML | | ✓ | ✓ |
| Rojas & Mukherjee (2006) | Journal of Computing in Civil Engineering | 29 | Univ Washington; Michigan Tech | USA | Web | Building general-purpose situational simulations | Representation; agent reasoning | | | |
| Eguchi (2016) | Robotics and Autonomous Systems | 28 | Bloomfield Coll | USA | Web | Educational robotics | | | | ✓ |
| Chi, VanLehn, Litman, & Jordan (2011) | User Modeling and User-adapted Interaction | 28 | Univ Pittsburgh et al. | USA | Web | Intelligent tutoring systems | RL | | ✓ | ✓ |
| Tsiriga and Virvou (2004) | User Modeling and User-adapted Interaction | 28 | Univ Piraeus | Greece | Web | Intelligent tutoring systems | ML; user modeling | | ✓ | ✓ |
| Tsai, Tseng, & Lin (2001) | Computational Science – Ivanović & Ho, 2019, Proceedings Pt 2 | 28 | Natl Chiao Tung Univ | Taiwan | Web | Adaptive learning environment | Fuzzy set; data mining; ML | | ✓ | ✓ |
| Cetintas, Si, Xin, and Hord (2009) | IEEE Transactions on Learning Technologies | 26 | Purdue Univ | USA | Web | Identifying off-task behaviors in intelligent tutoring systems | ObustRidge regression | | ✓ | ✓ |
| Nielsen, Ward, and Martin (2009) | Natural Language Engineering | 26 | Univ Colorado; Boulder Language Technol | USA | Web | Recognizing entailment in intelligent tutoring systems | NLP | | ✓ | ✓ |
| Rodrigues and Oliveira (2014) | Computers & Education | 25 | Inst Engn Polytech Porto | Portugal | Web | Formative assessment and monitoring of | | | | |
| students' progress | NLP | | ✓ | ✓ | | | | | | |
| Dolšak (2002) | Knowledge-Based Systems | 23 | Univ Maribor | Slovenia | Web | Expert system | | | | ✓ |
| Romero et al. (2013) | Expert Systems | 22 | Univ Cordoba | Spain | Web | Computer-based testing | Association rule mining; genetic programming | | ✓ | ✓ |
| Verdú, Verdú, Regueras , de Castro , and García (2012) | Expert Systems with Applications | 22 | Univ Valladolid | Spain | Web | Adaptive question sequencing system | Classification; genetic algorithms; fuzzy systems | | ✓ | ✓ |
| Tan, Ji, and Jin (2012) | IEEE Transactions on Education | 21 | Zhejiang Univ Technol et al. | China | Web | Parallel robot Instruction | ANN | ✓ | ✓ | ✓ |
| Tassopoulos and Beligiannis (2012) | Expert Systems with Applications | 21 | Univ W Greece, Hellenic Open Univ | Greece | NS | School timetabling problem | Particle swarm optimization | | ✓ | ✓ |
| Gaudioso, Montero, and Hernandez-Del-Olmo (2012) | Expert Systems with Applications | 21 | Univ Natl Educ Distancia; Inst Educ Secundaria | Spain | Web | Adaptive educational systems | ML | | ✓ | ✓ |
| Mozgovoy, Kakkonen, and Cosma (2010) | Journal of Educational Computing Research | 21 | Univ E Finland; Univ Aizu; PA Coll | Finland; Japan; Cyprus | Web | Plagiarism detection | NLP; information retrieval | | | |
| Luna, Romero, Romero, and Ventura (2015) | Applied Intelligence | 20 | Univ Cordoba; King Abdulaziz Univ | Spain; Saudi Arabia | Web | Mining rare class association rules | Association rules; genetic programming; evolutionary computation | | ✓ | ✓ |
| García, Schiaffino, and Amandi (2008) | Journal of Computer Assisted Learning | 20 | Univ Natl Ctr Provincia Buenos Aires | Argentina | Web | Detect students' | | | | |
| learning styles | Bayesian networks | | ✓ | ✓ | | | | | | |
| Piramuthu (2005) | IEEE Transactions on Education | 20 | Univ Florida | USA | Web | Intelligent tutoring systems | ML | | ✓ | ✓ |

Note: C: citation count; NS: No specified.

**Table 5**
Theories and novel models or frameworks proposed in the influential AIEd studies.

| | Theory | Author(s) and year |
|---|---|---|
| Existent theories or frameworks | Learning styles | García et al. (2007; 2008) |
| | Situated learning | Thomaz and Breazeal (2008) |
| | Bi-directional theory | Miyamoto et al. (1996) |
| | Collaborative learning | Calvo et al. (2010) |
| | Personalized learning | Cheung et al. (2003); Lan et al. (2014) |
| | Adapting learning theory | Xu and Wang (2006) |
| Proposed theoretical frameworks | Theory of movement-pattern perception with the basis of the bi-directional theory | Miyamoto et al. (1996) |
| Proposed models, systems, algorithms, or methodological frameworks | Personalization model in personalization for virtual learning environments | Xu and Wang (2006) |
| | Multi-agent framework for general-purpose situational simulations | Rojas & Mukherjee (2006) |
| | Framework for initializing student models in a web-based application | Tsiriga and Virvou (2004) |
| | Bayesian network modeling of learning styles | García et al. (2007; 2008) |
| | Dropout prediction approach for e-learning | Lykourentzou et al. (2009) |
| | A combinational incremental ensemble of classifiers for the prediction of learner performance in distance learning | Kotsiantis et al. (2010) |
| | Architecture for supporting collaborative writing on the cloud | Calvo et al. (2010) |
| | Analytical model for predicting and explaining reasons behind freshmen student attrition | Delen (2010) |
| | Adaptive algorithm with the basis of evolutionary computation for solving the timetabling problem of educational institutions | Beligiannis et al. (2008) |
| | A computational model to develop semantic web-driven systems | Bittencourt et al. (2009) |
| | A framework of machine learning-driven learning analytics | Lan et al. (2014) |
| | A computer program that aids the selection of learning management systems | Cavus (2010) |
| | A semiautomatic framework for producing domain concept maps and deriving domain ontologies | Zouaq & Nkambou (2008) |
| | Intelligent tutoring system for distance education | Cheung et al. (2003) |
| | Neural network-driven approach for predicting learners' mood | Moridis and Economides (2009) |
| | Computational gaming detector that automatically labels Intelligent Tutoring System data | Muldner et al. (2011) |
| | Two-phase fuzzy mining and learning approach for adaptive learning | Tsai et al. (2001) |
| | Machine learning algorithm to automatically detect learners' off-task behaviors | Cetintas et al. (2009) |
| | Formative assessment tool | Rodrigues and Oliveira (2014) |
| | Consultative rule-driven expert system for finite element mesh design | Dolšak (2002) |
| | Genetic fuzzy expert system to automatically classify questions | Verdú et al. (2012) |
| | Intelligent computer-aided instruction modeling for parallel robot instruction | Tan et al. (2012) |
| | Adaptive algorithm with the basis of particle swarm optimization to solve high school timetabling problem | Tassopoulos and Beligiannis (2012) |
| | Evolutionary algorithm to mine rare class association rules | Luna et al. (2015) |
| | A framework to capture both the dynamics of knowledge creation of learners | Piramuthu (2005) |

### 4.5. Top frequently used terms in highly cited AIEd

The top frequently used terms in the highly cited AIEd articles are shown in Table 8. The most frequently used one was "machine learning (appears in 13 articles, occupies 28.89%)," showing that the adoption of ML techniques in education context was a great concern among authors. Other frequently used terms included "artificial intelligence (12, 26.67%)," "tutoring system (12, 26.67%)," "intelligent tutoring system (8, 17.78%)," "neural network (6, 13.33%)," and "learning technique (6, 13.33%)."

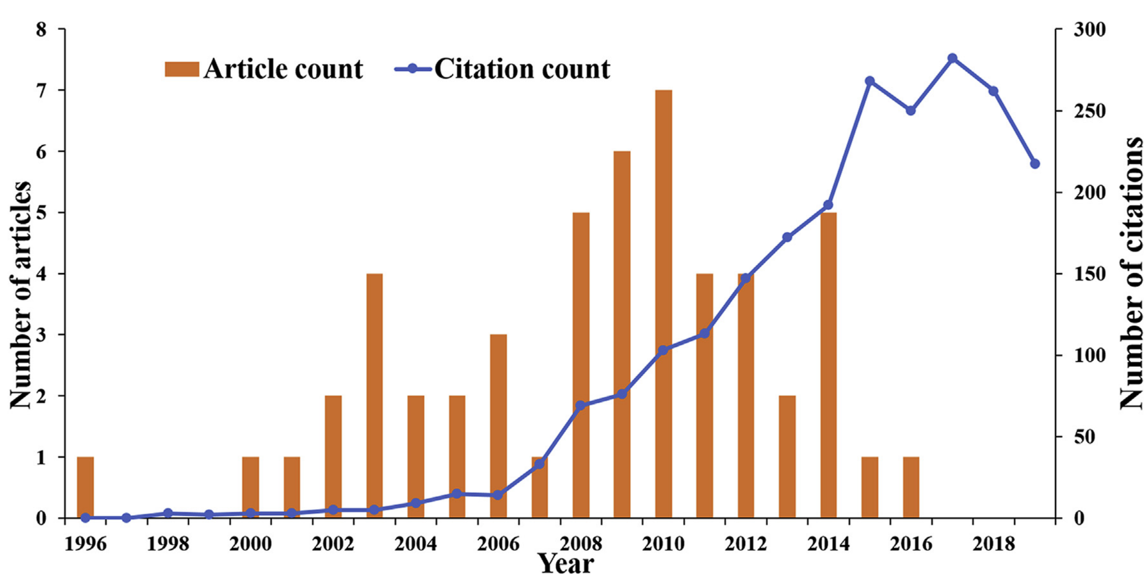



**Fig. 6.** Distributions of AIEd, MLEd, and DLEd over three periods.

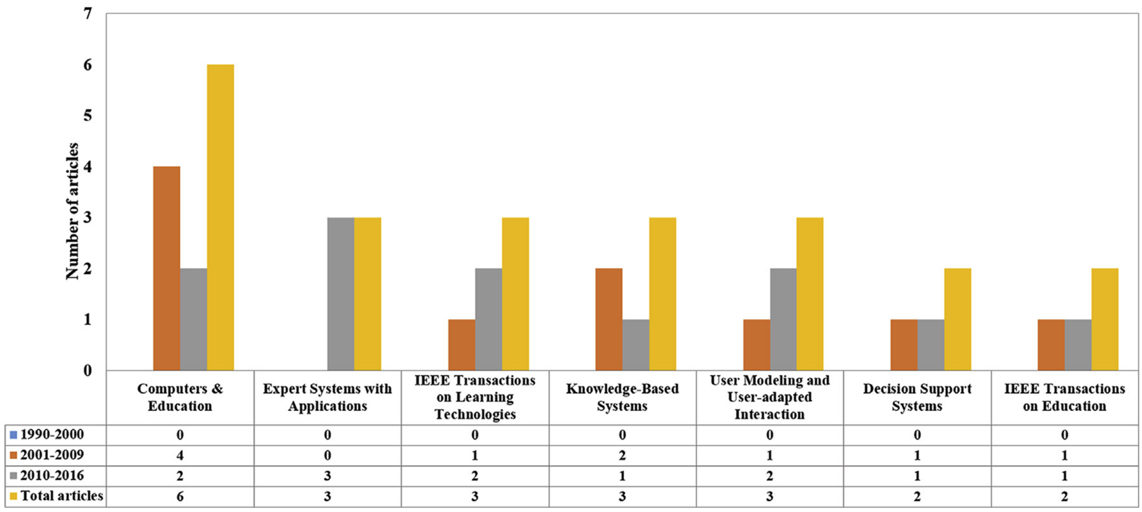



**Fig. 7.** Article count and citation count of AIEd.

**Table 6**
Top journals in publishing highly cited AIEd articles.

| Journals | A | TC | ACP | IF | 5-year IF | WoS Category | Position |
|---|---|---|---|---|---|---|---|
| Computers & Education | 6 | 571 | 95.17 | 5.627 | 5.902 | Computer Science, Interdisciplinary Applications; | 7 of 106, Q1 |
| Expert Systems with Applications | 3 | 64 | 21.33 | 4.292 | 4.577 | Computer Science, Artificial Intelligence | 24 of 134, Q1 |
| IEEE Transactions on Learning Technologies | 3 | 117 | 39.00 | 2.315 | 3.213 | Computer Science, Interdisciplinary Applications | 50 of 106, Q2 |
| Knowledge-Based Systems | 3 | 132 | 44.00 | 5.101 | 5.358 | Computer Science, Artificial Intelligence | 17 of 134, Q1 |
| User Modeling and User-adapted Interaction | 3 | 86 | 28.67 | 3.4 | 4.131 | Computer Science, Cybernetics | 4 of 23, Q1 |
| Decision Support Systems | 2 | 91 | 45.50 | 3.847 | 4.903 | Computer Science, Artificial Intelligence | 30 of 134, Q1 |
| IEEE Transactions on Education | 2 | 41 | 20.50 | 2.214 | 2.617 | Education, Scientific Disciplines | 15 of 41, Q2 |

Note: A: article count; TC: citation count; ACP: citations per article.

## 5. Discussions

This systematic literature review aimed to offer a synthesis of the best available evidence from 45 influential AIEd studies, examining the development, trends, and technologies adopted, as well as major research issues concerned by the AIEd community. In this section, we present the major findings by answering the four research questions, implications for future design and application of educational systems, the latest trends in AIEd research, limitations, and suggestions for future AIEd research.

### 5.1. (RQ1) How do broad and narrow AIEd studies distribute?

According to Fig. 6, there is an increase in interest in AIEd research since 2001, particularly during the period 2008–2014, as witnessed in Fig. 7. This indicates an increase in the importance of the application of AI technologies in education. Such a finding was consistent with Hinojo-Lucena et al. (2019), which found an increasing tendency of and a worldwide interest in AIEd. More importantly, from citation perspective, the citations of the studied 45 influential AIEd studies have grown consistently over time, indicating a continuingly increasing impact brought by these studies. In addition, among the 45 studies, only two

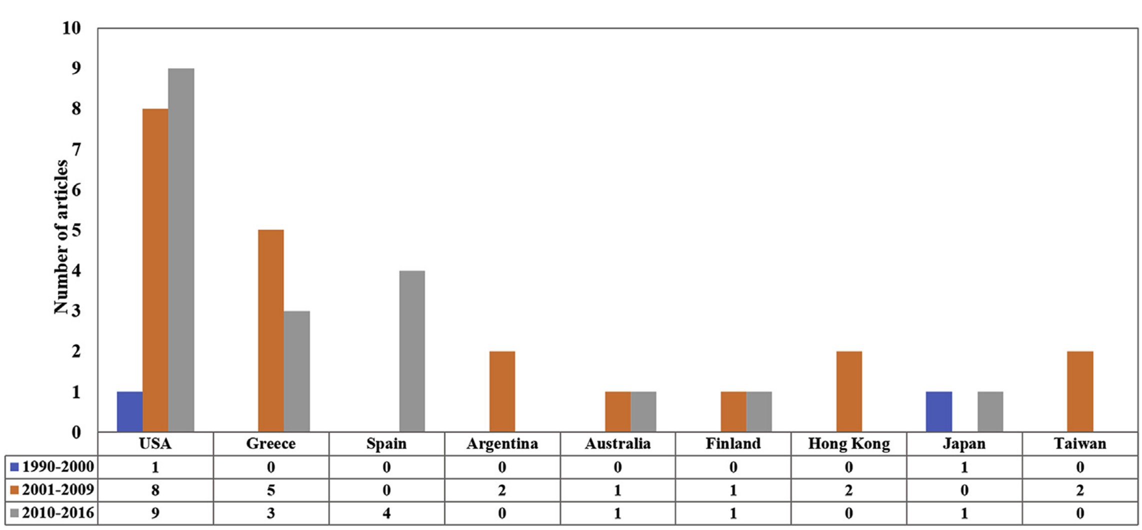



**Fig. 8.** Article distributions of the top journals for publishing influential AIED articles over three periods.

**Table 7**
Top countries/regions and institutions in publishing highly cited AIEd articles.

| C | A | TC | ACP | Institutions | A | TC | ACP |
|---|---|---|---|---|---|---|---|
| USA | 18 | 1079 | 59.94 | Stanford University | 3 | 224 | 74.67 |
| Greece | 8 | 364 | 45.50 | University of Patras | 3 | 125 | 41.67 |
| Spain | 4 | 85 | 21.25 | Arizona State University | 2 | 58 | 29.00 |
| Argentina | 2 | 208 | 104.00 | Carnegie Mellon University | 2 | 65 | 32.50 |
| Australia | 2 | 93 | 46.50 | Georgia Institute of Technology | 2 | 202 | 101.00 |
| Finland | 2 | 50 | 25.00 | Hellenic Open University | 2 | 83 | 41.50 |
| Hong Kong | 2 | 69 | 34.50 | Massachusetts Institute of Technology | 2 | 202 | 101.00 |
| Japan | 2 | 104 | 52.00 | University of Cordoba | 2 | 42 | 21.00 |
| Taiwan | 2 | 65 | 32.50 | National University of Central Buenos Aires | 2 | 208 | 104.00 |

Note: C: Countries/regions; A: article count; TC: citation count; ACP: citations per article.

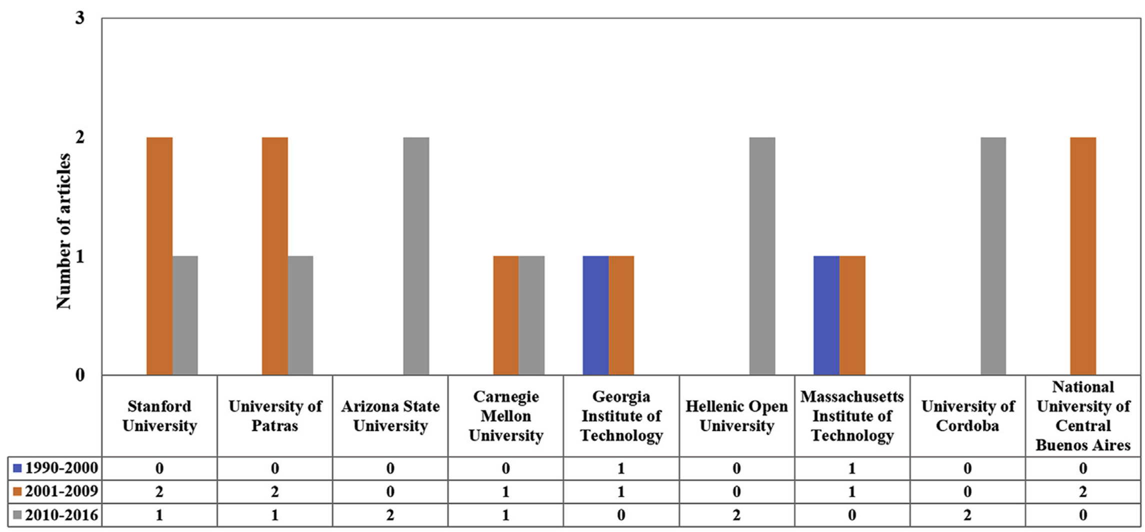



**Fig. 9.** Article distributions for highly cited countries/regions over three periods.

were found to be in relation to DLEd. As indicated by Khajah, Lindsey, and Mozer (2016), despite the broad application of DL technologies in other domains, they have been relatively underutilized in the education domain. Thus, the limitation of influential DLEd studies is understandable since DLEd is a younger area, and research on the application of DL technologies in educational contexts is under initial development. However, given the quick advance of deep learning, as reported by many studies (e.g., Peters, 2018; Tang, Peterson, & Pardos, 2016), DL's application in education is of dramatic potential. For example, as pointed by Tang et al. (2016), "recurrent neural networks and other deep learning techniques are impressively finding insights in many fields when it comes to large sets of often uncategorized data, and there are clear applications where education and learning can be improved, specifically where large amounts of student actions have already been collected (p. 4)." Therefore, it is expected that more influential studies concerning the application of DL technologies will increase in the near future. Hence, policymakers and practitioners are encouraged to propose a suitable strategy to allocate necessary resources so as to support future research and development of the application of AI technologies, particularly deep learning technologies in educational contexts.

### *5.2. (RQ2) What journals, institutions, and countries/regions contributed the most to the highly cited AIEd studies?*

According to Table 6, *Computers & Education* was found to be with the most highly cited AIEd articles. As a top journal in educational technology according to the Google Scholar,[3] *Computers & Education* has been found to be within the top journals in various subareas regarding education domains, for example, classroom dialogue (Song, Chen, Hao, Liu, & Lan, 2019) and augmented reality in science learning (Arici, Yildirim, Caliklar, & Yilmaz, 2019). In addition, many studies (e.g., Chen, Yu, Cheng, & Hao, 2019; Echeverria, 2019; Zawacki-Richter & Latchem, 2018) have targeted at the publications of *Computers & Education* to study theoretical perspectives and interpretations, research findings, and application practices concerning educational technology. Among the 45 influential AIEd studies, the top two all belonged to *Computers & Education*. The most influential one with a citation of 193 was conducted by Chen et al. (2010), in which they adopted a hierarchical linear model and multiple regressions to explore how web-based learning affected students' engagement and learning outcomes. The second influential study receiving a citation count of 188 was by García et al. (2007), in which they proposed the Bayesian networks model to detect learners' learning styles web-supported learning systems. Their study demonstrated the effectiveness of the proposed Bayesian model in detecting learners' learning styles. Figure A1 in Appendix shows the general distributions of AIEd-related articles in six top-tier educational technology journals during the period 1999–2019. As we can see, *Computers & Education* is in the first place with the number of articles as 63, far more than any other five journals, indicating currently, *Computers & Education* as the leading scientific journal welcoming studies concerning AIEd. In addition, regarding the rest of the journals listed in Table 6, none of them are closely related to educational technology, instead, most of which are in close relation to domains of computer science. By further referring to our findings concerning theories adopted and novel models or framework proposed in the influential AIEd studies (Table 5), it is found that there was a lack of educational theories being adopted in current AIEd studies, and most studies are concerned about proposing models, systems, or

3 https://scholar.google.com/citations?view_op=top_venues&hl=en&vq=eng_educationaltechnology.

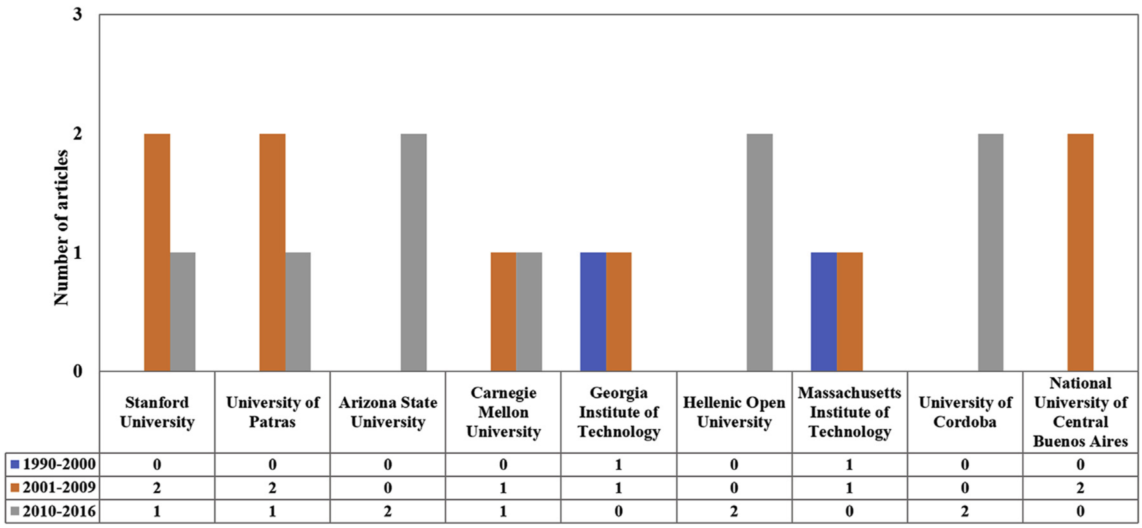



**Fig. 10.** Article distributions for highly cited institutions over three periods.

**Table 8**
Top frequently used terms in highly cited AIEd articles.

| Terms | F | % | Terms | F | % |
|---|---|---|---|---|---|
| machine learning | 13 | 28.89% | educational data mining | 5 | 11.11% |
| artificial intelligence | 12 | 26.67% | Bayesian network | 4 | 8.89% |
| tutoring system | 12 | 26.67% | different way | 4 | 8.89% |
| intelligent tutoring system | 8 | 17.78% | e-learning | 4 | 8.89% |
| Algorithm | 7 | 15.56% | experimental result | 4 | 8.89% |
| learning technique | 6 | 13.33% | learning process | 4 | 8.89% |
| neural network | 6 | 13.33% | natural language processing | 4 | 8.89% |

Note: F: frequency; %: the proportion of articles, meaning that the proportion of articles containing the term in the total studied articles.

methodological frameworks, while few of them are about proposing a novel theoretical framework. It is thus suggested that there is a need to incorporate the application of AI technologies with educational theories. In this way, AIEd studies with both concerns of technologies and educational theories tend more to be preferred by experts from both education domains and computer science domains and are possible to suits the scopes of journals regarding educational technology.

According to Song, Chen, Hao, Liu, & Lan (2019), the analysis of predominant institutions and countries/regions in a particular research field helps identify potential partners to conduct and share research. From the country/region perspective, according to Table 7, the USA has contributed far more influential AIEd studies. By further referring to the results of the top countries/regions contributing to the publication of AIEd-related articles during the period 1999–2019, as displayed in Figure A2 in Appendix, two kinds of findings are highlighted. On the one hand, there are several countries/regions contributing the most to both the AIEd research as well as the influential studies, particularly the USA, Spain, Taiwan, Australia, and China. On the other hand, some countries/regions such as Greece and Argentina have played a more important role in the publication of influential AIEd studies, as compared to the publication of AIEd research. In addition, some countries/regions such as England and Canada are active in the publication of AIEd research, while they have contributed less to the publication of influential AIEd studies. In this sense, it is advised that these countries/regions should attach more importance to either the productivity or the scientific impact and quality of their AIEd research. From an institution perspective, according to Table 7, *Stanford University* and *University of Patras* were found to be with the most influential AIEd studies. By further referring to the results of the top institutions contributing to the publication of AIEd-related articles during the period 1999–2019, as displayed in Figure A3 in Appendix, it is found that several institutions have contributed the most to both the AIEd research as well as the influential studies, particularly, *Stanford University*, *Carnegie Mellon University*, *Georgia Institute of Technology*, and *Massachusetts Institute of Technology*. However, some institutions, for example, *University of Patras* and *Arizona State University*, have contributed more to the publication of influential AIEd studies, as compared to the publication of AIEd research. Additionally, some institutions such as *University of Sydney* and *University of Pittsburgh* have contributed more to the publication of AIEd studies, as compared to the publication of influential AIEd studies. It is thus advised that these institutions should pay attention to either the productivity or the scientific impact and quality of their AIEd research.

### *5.3. (RQ3) What are the major issues addresses and AI technologies adopted in the highly cited AIEd studies?*

Since it is commonly accepted that terms or keywords analysis helps understand research foci in a particular period (Song, Chen, Hao, Liu, & Lan, 2019), answers to RQ3 also provide insights into major themes being concerned the most in the AIEd community. According to Tables 8 and 3 which depicts the types of AI technologies being adopted in each study, several important implications have been obtained.

Firstly, "machine learning" appeared the most frequently within the 45 highly cited AIEd articles with a percentage of 76%, which indicated that the adoption of ML techniques for educational purposes is a major concern among researchers in the AIEd research area. In addition to the top two influential studies mentioned in Section 5.2, there are many MLEd studies that are with a wide influence within academia. For example, a study by Thomaz and Breazeal (2008) with a citation count as 119 aimed to examine whether human-offered reward coincided with the conventional RL reward indication. They introduced an experimental platform based on a virtual RL robot and analyzed real-time human teaching behavior by examining untrained participants in teaching the robot to complete tasks. Their work highlighted the importance of comprehending the relationship between human-instructor and robot-learner so as to better propose approaches to help people in teaching and meanwhile enhance a robot's learning behavior. Besides, a study conducted by Pierre-Yves (2003) (receiving 116 citations) enabled a robot to express its emotions, by continuously controlling the age of a synthetic voice as well as the number of emotions expressed.

Secondly, there is a diversity in the use of terminology, with terms such as "tutoring system" and "intelligent tutoring system," developing from general to concrete area. According to Tsiriga and Virvou Tsiriga and Virvou (2004), ITSs serve as computer programs aiming to provide cost-effective instruction. A growing tendency in the adoption of computers to teach results in ITSs' development (Cetintas et al., 2009). ITSs are effective in increasing learners' learning engagement (Suchman, 1987) and learning outcomes (Koedinger, Anderson, Hadley, & Mark, 1997). Many researchers are striving to provide more ITSs by adopting ML technologies, and are receiving wide attention from academia. For example, Hämäläinen and Vinni (2006) provided general outlines concerning the classification of small data sets that were both numeric and categorical types and recommended variations of robust naïve Bayes classifiers that were able to deal with diverse variables and offer useful implications. Nielsen et al. (2009) indicated that highly effective discourse and instruction within ITSs could only be realized when systems were capable of recognizing detailed entailment relationships between learners' answers and the preferred conceptual understanding. However, currently, relatively fewer studies have been found to target on such issues.

Thirdly, neural network-related algorithms have been popularly adopted for educational purposes, with frequently used terms such as "neural network," "bayesian network," and "network." Such technologies have found to be applied to tackle various issues, for example, the detection of a student's learning style (García et al., 2007), dropout prediction in e-learning courses (Lykourentzou et al., 2009), student retention management (Delen, 2010), prediction of student's mood (Moridis & Economides, 2009), prediction of students' gaming behaviors (Muldner et al., 2011), improvement of teaching parallel robots (Tan et al., 2012), and detection of students' learning styles (García et al., 2008). Such algorithms as Bayesian networks (e.g., García et al., 2007; García et al., 2008; Muldner et al., 2011) and neural networks (e.g., Lin & Lu, 1995; Lykourentzou et al., 2009; Moridis & Economides, 2009; Pierre-Yves, 2003) have been commonly found to be adopted for educational purposes, while more advanced deep learning algorithms, for example, spiking neural networks (Kasabov, 2014) and artificial neural network (Tan et al., 2012) have seldom been found to arose great influence in educational domains. However, the application of such advanced algorithms as generative adversarial network and convolutional neural network are seldom adopted in educational domains, for example, generation of educational game levels (Park et al., 2019), and classification of cross-domain MOOC forum post (Wei, Lin, Yang, & Yu, 2017). However, such algorithms are effective in various domains, for example, medical (Nie et al., 2017), biology (Cao, Bhattacharya, Hou, & Cheng, 2016), transportation (Zhao, Zhou, Lu, & Fujita, 2019), and social media (Ma et al., 2016). Therefore, it is suggested that scholars in the area of AIEd are encouraged to devote more effort to applying advanced deep learning technologies to educational domains.

Besides, with a concern about the interactions between machines and natural languages, NLP technologies have also been found frequently adopted within the influential AIEd studies, particularly areas regarding assessment, for example, plagiarism detection (Mozgovoy et al., 2010), assignment grading (Burstein et al., 2004), and assessment of assignments and provision of automatic feedback (Rodrigues & Oliveira, 2014). Significant techniques that have been commonly incorporated into current computer-assisted assessment systems include statistical, latent semantic analysis, and NLP, within which NLP-driven systems are prevalent with the ability to conduct intelligent analyses by capturing the semantic meaning of unstructured texts (Rodrigues & Oliveira, 2014). Theoretically, NLP techniques have been adopted in educational domains for studies concerning reading, writing, and content knowledge teaching since the 1940s, 1960s, and the early 1970s, respectively (Burstein, 2009). In this sense, as compared to the history of NLP in educational research, there leaves much to be desired and achieved in making a breakthrough in such an area of application. For example, NLP technologies have great potential to provide and advance precision education and personalized education, particularly based on individual learners' portfolios.

In addition, how learners learn and perform during the process of learning is also a great concern by the AIEd scholars, with most of the relevant influential AIEd are concerned about adaptive or personalized learning (e.g., García et al., 2008; Gaudioso et al., 2012; Xu & Wang, 2006). For example, based on a developed personalization model, Xu and Wang (2006) proposed an intelligent decision-making agent for realizing personalization within personalized virtual learning environments (PVLE) with the use of intelligent agent' s autonomous, pre-active and proactive behaviors. The application of AI technologies in the evaluation of the learning process can be advanced in two directions, including the combination with electroencephalogram technologies to enhance effectiveness and efficiency, and the application into emotion-related issues such as attention recognition or emotion detection.

### 5.4. (RQ4) Theories and frameworks within the influential studies

Based the results presented in Table 5, several implications can be obtained. To begin with, educational theories have not been commonly adopted within the influential studies, with only several theories found. Such finding was similar to the study by Zawacki-Richter et al. (2019). First, a learning-style model classifies students according to the ways where they obtain and deal with information (García et al., 2007). García et al. (2008) presented a novel Bayesian model for detecting learners' learning styles. Second, situated learning theory concerns children's social world and its contribution to their development. Inspired by the situated learning theory, Thomaz and Breazeal ((2008) advocated an innovative idea that regarded ML issues to be the interaction between humans and machines. Third, bi-directional theory (Kawato, 1996; Meyer & Kornblum, 1993) provided a general computational framework for sensory-motor integration and generic representations derivation for various motor behaviors. Base on the bi-directional theory, Miyamoto et al. (1996) proposed a movement-pattern perception theory for motion capture and learning. Fourth, according to Lowry, Curtis, and Lowry (2004), collaborative writing is "an iterative and social process that involves a team focused on a common objective that negotiates, coordinates, and communicates during the creation of a common document (p. 72)." In the study by Calvo et al. (2010), they described a framework for supporting collaborative writing. In the study by Cheung et al. (2003), they described the design of SmartTutor, which integrated educational theories and AI into a tutoring system for the provision of personalized advices to learners according to their background and experience. In the study by Lan et al. (2014), they demonstrated that learners' concept knowledge, question–concept association, and intrinsic question difficulty were able to enhance ML-driven personalized learning. In addition, with reference to adapting learning theory, Xu and Wang (2006) developed personalized virtual learning environments to achieve learning effectiveness.

Furthermore, in terms of the newly proposed theoretical framework, only one relevant study was found, in which Miyamoto et al. (1996) proposed movement-pattern perception theory with the basis of bi-directional theory concerning the integration of sensory-motor. With the use of a toy task Kendama in the experiment, they demonstrated that the proposed theory was effective in capturing and learning motion. Additionally, the majority of the influential studies are found to be concerned with proposing novel models (e.g., Cetintas et al., 2009; Xu & Wang, 2006), systems (e.g., Cheung et al., 2003; Dolšak, 2002), algorithms (e.g., Luna et al., 2015; Tassopoulos & Beligiannis, 2012), and methodological frameworks (e.g., Calvo et al., 2010; Lowry et al., 2004) or the direct application of AI technologies in tackling educational issues (e.g., Hämäläinen & Vinni, 2006; Romero et al., 2013).

### 5.5. Implications about future design and application of educational systems

Based on the above discussions, some important implications are obtained, as elaborated in the following paragraphs, with foci on the application and theoretical perspectives.

From the perspective of the application, several potential research directions can be highlighted. First, as indicated in Table 3, most influential AIEd studies are concerned about the application of AI technologies in the contexts of online or web learning, while few concerned about the promotion of learning and teaching in physical contexts with the help of AI technologies. Thus, it is highly suggested that scholars in computer science and education should pay more attention to seek the potential of applying AI technologies in physical classroom settings for enhancing the learning and teaching process. Second, as within the influential studies, 11 are found to be related to ITSs; thus, in order to realize truly effective pedagogy using ITS, scholars are encouraged to spare more effort to recognize detailed entailment relationships between a learners' answer and the desired conceptual understanding. Third, although NLP technologies have long been adopted in tackling educational issues, through our analysis, a limited breakthrough has been brought in such an area. As a powerful technique for dealing with unstructured text data, NLP techniques are effective in various areas for achieving personalized purposes such as precision medicine. Thus, it is suggested that scholars seek NLP's potential in promoting precision or personalized education. Furthermore, as the learning process is found to be highly concerned within the AIEd community, we here provide two potential directions for the development of AI in tackling issues regarding learning process, that is, to combine with biomedical detection and imaging technologies such as electroencephalogram, and to target at issues related to learners' affection and emotion.

In addition, according to the analysis of the AI technologies adopted in the influential studies, it is found that traditional AI technologies, such as NLP and ML methods such as neural network and Bayesian networks, are the most commonly adopted in educational contexts. However, advanced deep learning algorithms such as generative adversarial networks and deep neural networks have seldom been found. Although such findings are understandable since deep learning is a relatively younger research area as compared to AI and ML, however, given the rapid advancement and effectiveness of DL, the application of DL in education domains should be enhanced and paid more attention to. Specifically, we here provide several suggestions concerning applying DL technologies into resolving issues concerning teaching and learning. First, scholars are advised to seek the potential of applying deep neural networks in the context of affect and behavior modeling to build detectors of student affective states and behaviors. Second, educators and researchers can pay more attention to monitor learners' attendance as well as their careful and careless states through face detection and recognition in videos and images. Third, it can be significant to facilitate the automatic marking of learners' reading comprehension by using long short-term memory (LSTM) neural network. Fourth, deep neural networks is potential for game-based learning stealth assessment. Fifth, scholars are advised to facilitate assignment grading by using a convolutional neural network. Sixth, learning behavior analysis and dropout rate prediction can be facilitated by the use of the LSTM neural network. Seventh, it is useful to enhance the prediction of question quality with the use of recurrent neural networks. Furthermore, educators and researchers are advised to seek the potential of predicting learners' academic persistence by using a recursive neural network and LSMT algorithm. In addition, it is suggested to facilitate the prediction of learners' performance with the use of generative adversarial networks.

In the application of AI technologies to the analysis of educational data, issues concerning ethical and privacy are worth noting. The application of AI technologies commonly requires large amounts of data, involving confidential information about students and faculty. Thus, issues concerning privacy and data protection may involve, which should be considered seriously when developing research and practical activities regarding AIEd.

From a theoretical perspective, based on the discussion in Section 5.2 and 5.4, it is found that many influential AIEd studies are published by computer science journals, and most of the 45 studies are related to the proposal of new models, systems, algorithms, or methodological frameworks. However, educational theories have seldom been adopted, let along the proposal of novel theories or theoretical frameworks. It is thus suggested that there is a need for a close combination of AI technologies with educational theories. Such studies can truly promote the development of AIEd as an important and promising subfield in the area of educational technology.

### 5.6. Latest trends in AIEd research

We also provide examples of the most up-to-date achievements of AIEd research. To begin with, there is a trend toward emotion detection in game-based learning. For example, using ML techniques, Ninaus et al. (2019) aimed at automatically detecting emotions from facial data. Their study demonstrated an increase in positive and negative emotions during the process of game-based learning. Second, there is an increasing interest in adopting ML algorithms to deal with issues concerning adaptive learning systems. For example, with the integration of item response theory and ML algorithms, Pliakos et al. (2019) presented an innovative system for dealing with the inability problem of providing learning materials to new users of adaptive learning systems. Experimental results indicated that the proposed system was effective in predicting learning materials for new users. Third, some scholars are concerned about student retention prediction using ML techniques. For example, based on several developed statistics, Gray and Perkins (2019) proposed a retention prediction model using ML algorithms. Their work has proven the effectiveness of the proposed model in supporting learners' learning. Fourth, there is also a growing trend in the study of teacher responding tool. For example, driven by a common problem that it was difficult for instructors to give an immediate and effective response to learners' mathematical ideas in practice, Bywater, Chiu, Hong, and Sankaranarayanan (2019) introduced teacher responding tool with the use of NLP techniques and evaluated its usefulness in providing automatic and effective respond recommendations to instructors. Moreover, the role of AI has been widely debated and has received increasing attention from academia. For example, in the study by Cukurova, Kent, & Luckin (2019), they put forward an idea of human intelligence argumentation supported by AI techniques. Such an idea was illustrated using a case study of debate tutoring where prediction and classification algorithms were used to enhance the decision-making of tutors in providing reflections and feedback. Furthermore, a growing trend has been noted in the study of how to use robots to support learners with intellectual disabilities to learn. For example, using eye-gaze as outcome measurement, Hughes-Roberts et al. (2018) explored whether robots could meet the needs of learners with intellectual disabilities to enhance individual and personalized learning in terms of engagement and achievement. In addition, the use of AI techniques in supporting dialogue analysis has also become popular among AIEd researchers. For example, using Bakhtin's notion of speech genres as a theoretical framework, Sullivan & Keith (2019) investigated the potential of adopting NLP techniques to enhance dialogue analysis during small-group learning.

### 5.7. Limitations and areas for future research

A systematic overview of AIEd research is presented by the use of 45 influential studies. However, this study is subject to certain methodological limitations. First, if an article was not published in a computer science or education-related journal indexed by SCI or SSCI database, then it would be excluded. There are AIEd-related journals, for example, *International Journal of Artificial Intelligence in Education*, are not included in the SCI or SSCI database. Thus, it is likely that some AIEd journal articles were not included in our review. Second, only research articles were considered, while conference papers and editorials were not included. However, since AIEd is of interdisciplinary nature with scholars from particularly computer science and education areas, a number of studies might be published as conference papers that are also of great influence. Thus, it is also possible that some influential AIEd conference papers are not included in the analysis. Another limitation is that we only included articles with a citation count not less than 20. Since in addition to the quality itself, it is also a matter of time for an academic publication to get citations.

Within the 45 studied studies, the most latest one was published in the year 2016. That is to say, the latest AIEd studies being published during 2017–2019 were not considered, perhaps not because of quality but without enough time to get cited. Therefore, some of the results of this study may fail to reflect the most recent development. However, as we focus on the most influential AIEd research until now, this study has lived up to the expectation to reflect what the AIEd community concern the most. In addition, as the current WoS does not self-citations, it is potential that some of the studied articles may not be truly influential.

Considering the above limitations, directions for future research may include the following aspects. First, to include recent AIEd studies, future research can consider reducing the inclusion criteria, citation count. In this way, more recent studies can be included. Second, a bibliometric analysis based on more comprehensive AIEd research data (e.g., AIEd research articles and conference papers retrieved from different databases with no restriction of citation count) can be considered to bring a general picture of the status and structure of the whole AIEd landscape. Third, with large-scale AIEd literature data, the use of text mining and NLP techniques like structural topic modeling can provide a more general understanding of AIEd research. In addition, future research can consider examining AIEd research in individual countries/regions to identify important institutions and scholars from a national perspective.

## 6. Conclusions

This study provides a systematic review of AIEd research. It is worth noting the practical application of this study as it offers essential information to better understand the most concerning issues in AIEd. Results obtained are beneficial to AIEd practitioners and scholars. First, findings obtained assist newcomers to AIEd in finding theories, tools, and techniques that are commonly adopted by influential AIEd studies. Second, this study helps scholars recognize important institutions and countries/regions that have made significant contributions to AIEd research to further seek collaborations or scientific exchanges. In addition, this study helps researchers in understanding important topics and future directions concerning AIEd research.

## Declaration of competing interest

As the research does not involve human participants directly, there is no need to seek ethical approval from a research review committee in the authors' affiliations. The data can be obtained by sending emails to the corresponding author. The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Acknowledgements

The research described in this paper has been fully supported by the Research Seed Fund (102367), the Teaching Development Grant (102489), and LEO Dr David P. Chan Institute of Data Science, Lingnan University, Hong Kong.

## Appendix

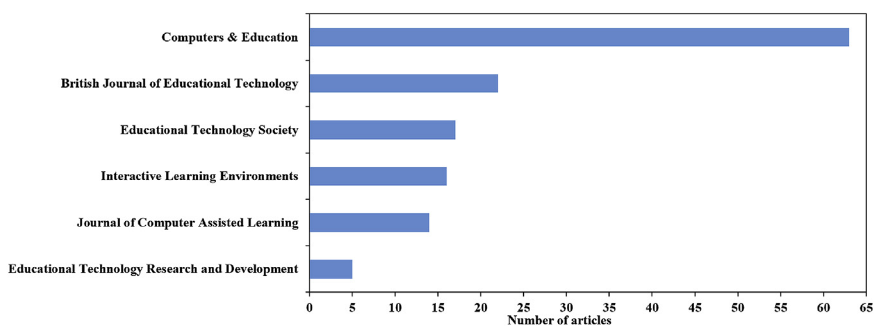



**Fig. A1.** Number of AIEd articles of six SSCI journals in the area of educational technology during the period 1990–2019

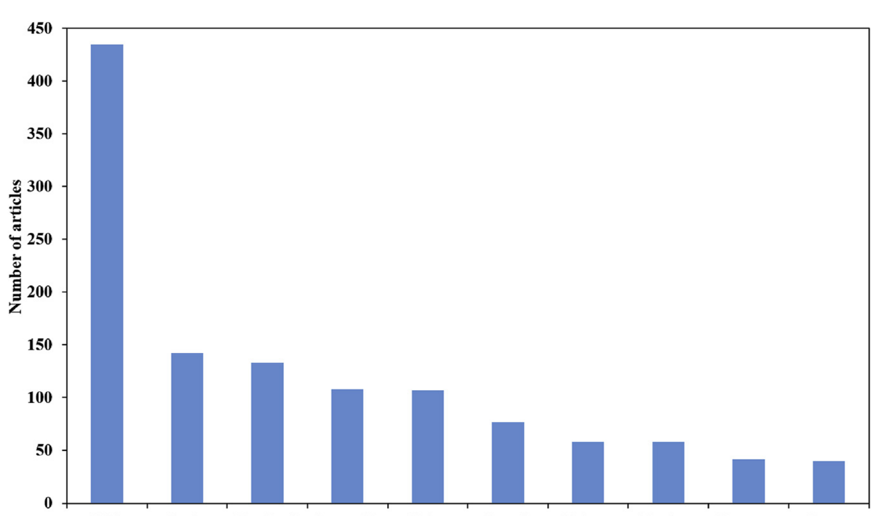



**Fig. A2.** Number of AIEd articles of top ten countries/regions during the period 1990–2019 (Note: Data were collected from WoS)

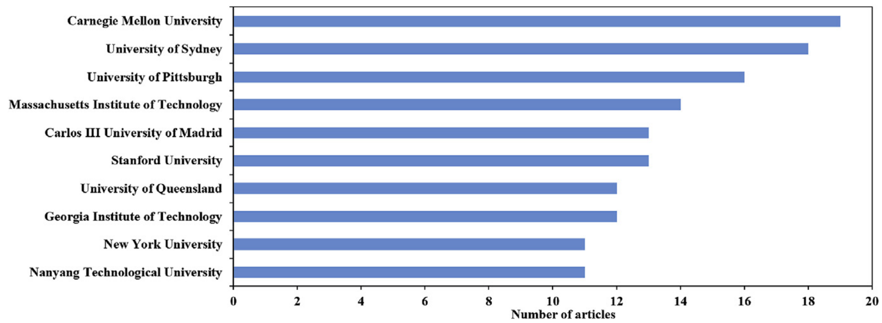



**Fig. A3.** Number of AIEd articles of top ten institutions during the period 1990–2019 (Note: Data were collected from WoS)

## References


Abramo, G., & D'Angelo, C. A. (2011). Evaluating research: From informed peer review to bibliometrics. *Scientometrics, 87*(3), 499–514.

Aizenberg, I., Aizenberg, N. N., & Vandewalle, J. P. (2013). *Multi-valued and universal binary neurons: Theory, learning and applications*. Springer Science & Business Media.

Arden, S. V., Pentimonti, J. M., Cooray, R., & Jackson, S. (2018). A categorical content analysis of highly cited literature related to trends and issues in special education. *Journal of Learning Disabilities, 51*(6), 589–599.

Arici, F., Yildirim, P., Caliklar, Ş., & Yilmaz, R. M. (2019). Research trends in the use of augmented reality in science education: Content and bibliometric mapping analysis. *Computers & Education, 142*, 103647.

Baepler, P., & Murdoch, C. J. (2010). Academic analytics and data mining in higher education. *International Journal for the Scholarship of Teaching & Learning, 4*(2), 1–9.

Baker, T., Smith, L., & Anissa, N. (2019). *Educ-AI-tion rebooted? Exploring the future of artificial intelligence in schools and colleges*.

Baker, R. S., & Yacef, K. (2009). The state of educational data mining in 2009: A review and future visions. *JEDM| Journal of Educational Data Mining, 1*(1), 3–17.

Baltussen, A., & Kindler, C. H. (2004). Citation classics in anesthetic journals. *Anesthesia & Analgesia, 98*(2), 443–451.

Bayne, S. (2015). Teacherbot: Interventions in automated teaching. *Teaching in Higher Education, 20*(4), 455–467.

Beligiannis, G. N., Moschopoulos, C. N., Kaperonis, G. P., & Likothanassis, S. D. (2008). Applying evolutionary computation to the school timetabling problem: The Greek case. *Computers & Operations Research, 35*(4), 1265–1280.

Bengio, Y., Courville, A., & Vincent, P. (2013). Representation learning: A review and new perspectives. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 35*(8), 1798–1828.

Bengio, Y., Lee, D.-H., Bornschein, J., Mesnard, T., & Lin, Z. (2015). *Towards biologically plausible deep learning. arXiv preprint arXiv:1502.04156*.

Biswas, G., Leelawong, K., Schwartz, D., Vye, N., & The Teachable Agents Group at Vanderbilt. (2005). Learning by teaching: A new agent paradigm for educational software. *Applied Artificial Intelligence, 19*(3–4), 363–392.

Bittencourt, I. I., Costa, E., Silva, M., & Soares, E. (2009). A computational model for developing semantic web-based educational systems. *Knowledge-Based Systems, 22*(4), 302–315.

Bostrom, N. (2006). *AI set to exceed human brain power*. CNN.

Botrel, L., Holz, E., & Kübler, A. (2015). Brain painting V2: Evaluation of P300-based brain-computer interface for creative expression by an end-user following the user-centered design. *Brain-Computer Interfaces, 2*(2–3), 135–149.

Brace, W. (1992). Quality assessment of library and information science school faculties. *Education for Information, 10*(2), 115–123.

Brynjolfsson, E., & Mitchell, T. (2017). What can machine learning do? Workforce implications. *Science, 358*(6370), 1530–1534.

Burstein, J. (2009). Opportunities for natural language processing research in education. In *International conference on intelligent text processing and computational linguistics* (pp. 6–27). Springer.

Burstein, J., Chodorow, M., & Leacock, C. (2004). Automated essay evaluation: The Criterion online writing service. *AI Magazine, 25*(3), 27-27.

Bywater, J. P., Chiu, J. L., Hong, J., & Sankaranarayanan, V. (2019). The Teacher Responding Tool: Scaffolding the teacher practice of responding to student ideas in mathematics classrooms. *Computers & Education, 139*, 16–30.

Calvo, R. A., O'Rourke, S. T., Jones, J., Yacef, K., & Reimann, P. (2010). Collaborative writing support tools on the cloud. *IEEE Transactions on Learning Technologies, 4*(1), 88–97.

Cao, R., Bhattacharya, D., Hou, J., & Cheng, J. (2016). DeepQA: Improving the estimation of single protein model quality with deep belief networks. *BMC Bioinformatics, 17*(1), 495.

Cavus, N. (2010). The evaluation of Learning Management Systems using an artificial intelligence fuzzy logic algorithm. *Advances in Engineering Software, 41*(2), 248–254.

Cetintas, S., Si, L., Xin, Y. P. P., & Hord, C. (2009). Automatic detection of off-task behaviors in intelligent tutoring systems with machine learning techniques. *IEEE Transactions on Learning Technologies, 3*(3), 228–236.

Chan, K. S., & Zary, N. (2019). Applications and challenges of implementing artificial intelligence in medical education: Integrative review. *JMIR medical education, 5*(1), Article e13930.

Chen, P.-S. D., Lambert, A. D., & Guidry, K. R. (2010). Engaging online learners: The impact of Web-based learning technology on college student engagement. *Computers & Education, 54*(4), 1222–1232.

Chen, X., Xie, H., Wang, F. L., Liu, Z., Xu, J., & Hao, T. (2018). A bibliometric analysis of natural language processing in medical research. *BMC Medical Informatics and Decision Making, 18*(1), 1–14.

Chen, X., Yu, G., Cheng, G., & Hao, T. (2019). Research topics, author profiles, and collaboration networks in the top-ranked journal on educational technology over the past 40 years: a bibliometric analysis. *Journal of Computers in Education, 6*(4), 563–585.

Chen, X., Zou, D., Cheng, G., & Xie, H. (2020). Detecting latent topics and trends in educational technologies over four decades using structural topic modeling: A retrospective of all volumes of computer & education. Computers & Education. *Computers & Education, 151*, Article 103855.

Chen, X., Zou, D., & Xie, H. (2020). Fifty years of British Journal of Educational Technology: A topic modeling based bibliometric perspective. *British Journal of Educational Technology, 51*(3), 692–708.

Cheung, B., Hui, L., Zhang, J., & Yiu, S.-M. (2003). SmartTutor: An intelligent tutoring system in web-based adult education. *Journal of Systems and Software, 68*(1), 11–25.

Chi, M., VanLehn, K., Litman, D., & Jordan, P. (2011). Empirically evaluating the application of reinforcement learning to the induction of effective and adaptive pedagogical strategies. User Modeling and User-Adapted. *Interaction, 21*(1–2), 137–180.

Choudhri, A. F., Siddiqui, A., Khan, N. R., & Cohen, H. L. (2015). Understanding bibliometric parameters and analysis. *RadioGraphics, 35*(3), 736–746.

Chuang, K.-Y., Wang, M.-H., & Ho, Y.-S. (2011). High-impact papers presented in the subject category of water resources in the essential science indicators database of the institute for scientific information. *Scientometrics, 87*(3), 551–562.

Cobo, M. J., Martínez, M. A., Gutiérrez-Salcedo, M., Fujita, H., & Herrera-Viedma, E. (2015). 25years at knowledge-based systems: A bibliometric analysis. *Knowledge-Based Systems, 80*, 3–13.

Confraria, H., Blanckenberg, J., & Swart, C. (2018). The characteristics of highly cited researchers in Africa. *Research Evaluation, 27*(3), 222–237.

Cukurova, M., Kent, C., & Luckin, R. (2019). Artificial intelligence and multimodal data in the service of human decision-making: A case study in debate tutoring. *British Journal of Educational Technology, 50*(6), 3032–3046.

Dechter, R. (1986). *Learning while searching in constraint-satisfaction problems*. University of California, Computer Science Department, Cognitive Systems ….

Delen, D. (2010). A comparative analysis of machine learning techniques for student retention management. *Decision Support Systems, 49*(4), 498–506.

Della Ventura, M. (2017). Creating inspiring learning environments by means of digital technologies: A case study of the effectiveness of WhatsApp in music education. In E-Learning, & E-Education (Eds.), *Online training* (pp. 36–45). Springer.

Della Ventura, M. (2018). Twitter as a music education tool to enhance the learning process: Conversation analysis. In *New media for educational change* (pp. 81–88). Springer.

Deng, L., & Yu, D. (2014). Deep learning: Methods and applications. *Foundations and Trends® in Signal Processing, 7*(3–4), 197–387.

Dolšak, B. (2002). Finite element mesh design expert system. In *Applications and innovations in intelligent systems IX* (pp. 3–14). Springer.

Echeverria, A. (2019). The impact of computers & education measured beyond traditional bibliographical metrics Echeverria, alejandro; nussbaum, miguel; albers, casper J.; heller, rachelle S.; Tsai, chin-chung; van braak, johan. *Computers & Education, 140*, 103592.

Eguchi, A. (2016). RoboCupJunior for promoting STEM education, 21st century skills, and technological advancement through robotics competition. *Robotics and Autonomous Systems, 75*, 692–699.

Florea, A. M., & Radu, S. (2019). Artificial intelligence and education. In *2019 22nd international conference on control systems and computer science (CSCS)* (pp. 381–382). IEEE.

Fu, H.-Z., Wang, M.-H., & Ho, Y.-S. (2012). The most frequently cited adsorption research articles in the Science Citation Index (Expanded). *Journal of Colloid and Interface Science, 379*(1), 148–156.
García, P., Amandi, A., Schiaffino, S., & Campo, M. (2007). Evaluating Bayesian networks' precision for detecting students' learning styles. *Computers & Education, 49*(3), 794–808.
García, P., Schiaffino, S., & Amandi, A. (2008). An enhanced Bayesian model to detect students' learning styles in Web-based courses. *Journal of Computer Assisted Learning, 24*(4), 305–315.
Garfield, E., & Merton, R. K. (1979). *Citation indexing: Its theory and application in science, technology and humanities* (Vol. 8). New York: Wiley.
Garg, T. (2020). Artificial intelligence in medical education. *The American Journal of Medicine, 133*(2), e68.
Gaudioso, E., Montero, M., & Hernandez-Del-Olmo, F. (2012). Supporting teachers in adaptive educational systems through predictive models: A proof of concept. *Expert Systems with Applications, 39*(1), 621–625.
Geng, Y., Chen, W., Liu, Z., Chiu, A. S., Han, W., Liu, Z., et al. (2017). A bibliometric review: Energy consumption and greenhouse gas emissions in the residential sector. *Journal of Cleaner Production, 159*, 301–316.
Godin, B. (2006). On the origins of bibliometrics. *Scientometrics, 68*(1), 109–133.
Gomez, F. J., & Schmidhuber, J. (2005). Co-evolving recurrent neurons learn deep memory POMDPs. In *Proceedings of the 7th annual conference on Genetic and evolutionary computation* (pp. 491–498). ACM.
Gori, M. (2017). *Machine learning: A constraint-based approach*. Morgan Kaufmann.
Gray, C. C., & Perkins, D. (2019). Utilizing early engagement and machine learning to predict student outcomes. *Computers & Education, 131*, 22–32.
Hallinger, P., & Chatpinyakoop, C. (2019). A bibliometric review of research on higher education for sustainable development, 1998–2018. *Sustainability, 11*(8), 2401.
Hämäläinen, W., & Vinni, M. (2006). Comparison of machine learning methods for intelligent tutoring systems. In *International conference on intelligent tutoring systems* (pp. 525–534). Springer.
Han, L. (2018). Analysis of new advances in the application of artificial intelligence to education. In *2018 3rd international conference on education, E-learning and management technology (EEMT 2018)*. Atlantis Press.
Hassabis, D., Kumaran, D., Summerfield, C., & Botvinick, M. (2017). Neuroscience-inspired artificial intelligence. *Neuron, 95*(2), 245–258.
Heffernan, N. T., & Heffernan, C. L. (2014). The ASSISTments Ecosystem: Building a platform that brings scientists and teachers together for minimally invasive research on human learning and teaching. *International Journal of Artificial Intelligence in Education, 24*(4), 470–497.
Heradio, R., de la Torre, L., Galan, D., Cabrerizo, F. J., Herrera-Viedma, E., & Dormido, S. (2016). Virtual and remote labs in education: A bibliometric analysis. *Computers & Education, 98*, 14–38.
Hinojo-Lucena, F.-J., Aznar-Díaz, I., Cáceres-Reche, M.-P., & Romero-Rodríguez, J.-M. (2019). Artificial intelligence in higher education: A bibliometric study on its impact in the scientific literature. *Education Sciences, 9*(1), 51.
Ho, Y.-S., & Hartley, J. (2017). Highly cited publications in world war II: A bibliometric analysis. *Scientometrics, 110*(2), 1065–1075.
Hughes-Roberts, T., Brown, D., Standen, P., Desideri, L., Negrini, M., Rouame, A., et al. (2018). Examining engagement and achievement in learners with individual needs through robotic-based teaching sessions. *British Journal of Educational Technology, 50*(5), 2736–2750.
Humble, N., & Mozelius, P. (2019). Artificial intelligence in education-a promise, a threat or a hype?. In *European conference on the impact of artificial intelligence and robotics 2019 (ECIAIR 2019)* (pp. 149–156). Oxford, UK: Academic Conferences and Publishing International Limited.
Ivanović, L., & Ho, Y.-S. (2019). Highly cited articles in the education and educational research category in the social science citation index: A bibliometric analysis. *Educational Review, 71*(3), 277–286.
Jordan, M. I., & Mitchell, T. M. (2015). Machine learning: Trends, perspectives, and prospects. *Science, 349*(6245), 255–260.
Kaplan, A., & Haenlein, M. (2019). Siri, siri, in my hand: Who's the fairest in the land? On the interpretations, illustrations, and implications of artificial intelligence. *Business Horizons, 62*(1), 15–25.
Karmani, P., Chandio, A. A., Korejo, I. A., & Chandio, M. S. (2018). October). A review of machine learning for healthcare informatics specifically tuberculosis disease diagnostics. In *International conference on intelligent technologies and applications* (pp. 50–61). Singapore: Springer.
Kasabov, N. K. (2014). NeuCube: A spiking neural network architecture for mapping, learning and understanding of spatio-temporal brain data. *Neural Networks, 52*, 62–76.
Kawato, M. (1996). *Bidirectional theory approach to integration*.
Khajah, M., Lindsey, R. V., & Mozer, M. C. (2016). *How deep is knowledge tracing?*. arXiv preprint arXiv:1604.02416.
Knapp, S. (2006). *Artificial intelligence: Past, present, and future*. Vox of Dartmouth.
Koedinger, K. R., Anderson, J. R., Hadley, W. H., & Mark, M. A. (1997). *Intelligent tutoring goes to school in the big city*.
Koedinger, K. R., & Corbett, A. (2006). *Cognitive tutors: Technology bringing learning sciences to the classroom: Na*.
Kohavi, R., & Provost, F. (1998). Glossary of terms journal of machine learning. In *Obtido*.
Kotsiantis, S., Patriarcheas, K., & Xenos, M. (2010). A combinational incremental ensemble of classifiers as a technique for predicting students' performance in distance education. *Knowledge-Based Systems, 23*(6), 529–535.
Kotsiantis, S. B. (2012). Use of machine learning techniques for educational proposes: a decision support system for forecasting students' grades. *Artificial Intelligence Review, 37*(4), 331–344.
Kotsiantis, S. B., Pierrakeas, C. J., & Pintelas, P. E. (2003, September). Preventing student dropout in distance learning using machine learning techniques. In *International conference on knowledge-based and intelligent information and engineering systems* (pp. 267–274). Berlin, Heidelberg: Springer.
Lai, C. L. (2019). Trends of mobile learning: A review of the top 100 highly cited papers. *British Journal of Educational Technology, 51*(3), 721–742.
Laird, T. F. N., Shoup, R., Kuh, G. D., & Schwarz, M. J. (2008). The effects of discipline on deep approaches to student learning and college outcomes. *Research in Higher Education, 49*(6), 469–494.
Lan, A. S., Waters, A. E., Studer, C., & Baraniuk, R. G. (2014). Sparse factor analysis for learning and content analytics. *The Journal of Machine Learning Research, 15*(1), 1959–2008.
LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature, 521*(7553), 436.
Levitt, J., & Thelwall, M. (2008). The most highly cited Library and Information Science articles: Interdisciplinarity, first authors and citation patterns. *Scientometrics, 78*(1), 45–67.
Liao, H., Tang, M., Li, Z., & Lev, B. (2019). Bibliometric analysis for highly cited papers in operations research and management science from 2008 to 2017 based on essential science indicators. *Omega, 88*, 223–236.
Lin, C.-T., & Lu, Y.-C. (1995). A neural fuzzy system with linguistic teaching signals. *IEEE Transactions on Fuzzy Systems, 3*(2), 169–189.
Lowry, P. B., Curtis, A., & Lowry, M. R. (2004). Building a taxonomy and nomenclature of collaborative writing to improve interdisciplinary research and practice. *The Journal of Business Communication, 41*(1), 66–99, 1973.
Luckin, R., Holmes, W., Griffiths, M., & Forcier, L. B. (2016). *Intelligence unleashed: An argument for AI in education*.
Luna, J. M., Romero, C., Romero, J. R., & Ventura, S. (2015). An evolutionary algorithm for the discovery of rare class association rules in learning management systems. *Applied Intelligence, 42*(3), 501–513.
Lykourentzou, I., Giannoukos, I., Nikolopoulos, V., Mpardis, G., & Loumos, V. (2009). Dropout prediction in e-learning courses through the combination of machine learning techniques. *Computers & Education, 53*(3), 950–965.
Ma, J., Gao, W., Mitra, P., Kwon, S., Jansen, B. J., Wong, K.-F., et al. (2016). *Detecting rumors from microblogs with recurrent neural networks*.
Marblestone, A. H., Wayne, G., & Kording, K. P. (2016). Toward an integration of deep learning and neuroscience. *Frontiers in Computational Neuroscience, 10*, 94.
Martinez, M., & Sá, C. (2019). Highly cited in the south: International collaboration and research recognition among Brazil's highly cited researchers. *Journal of Studies in International Education, 24*(1), 39–58.
Martí-Parreño, J., Méndez-Ibáñez, E., & Alonso-Arroyo, A. (2016). The use of gamification in education: A bibliometric and text mining analysis. *Journal of Computer Assisted Learning, 32*(6), 663–676.
McHugh, M. L. (2012). Interrater reliability: The kappa statistic. *Biochemia Medica: Biochemia Medica, 22*(3), 276–282.
Merceron, A., & Yacef, K. (2004). Mining student data captured from a web-based tutoring tool: Initial exploration and results. *Journal of Interactive Learning Research, 15*(4), 319–346.
Meyer, D. E., & Kornblum, S. (1993). *Optimization and learning in neural networks for formation and control of coordinated movement*.
Ming, H. W., Hui, Z. F., & Yuh, S. H. (2011). Comparison of universities' scientific performance using bibliometric indicators. *Malaysian Journal of Library & Information Science, 16*(2), 1–19.
Mitchell, T. (1997). *Machine learning*. WBC/McGraw-Hill, Boston: MA.
Miyamoto, H., Schaal, S., Gandolfo, F., Gomi, H., Koike, Y., Osu, R., et al. (1996). A kendama learning robot based on bi-directional theory. *Neural Networks, 9*(8), 1281–1302.
Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *Annals of Internal Medicine, 151*(4), 264–269.
Moridis, C. N., & Economides, A. A. (2009). Prediction of student's mood during an online test using formula-based and neural network-based method. *Computers & Education, 53*(3), 644–652.
Mostow, J., & Beck, J. (2006). Some useful tactics to modify, map and mine data from intelligent tutors. *Natural Language Engineering, 12*(2), 195–208.
Mozgovoy, M., Kakkonen, T., & Cosma, G. (2010). Automatic student plagiarism detection: Future perspectives. *Journal of Educational Computing Research, 43*(4), 511–531.
Muldner, K., Burleson, W., Van de Sande, B., & VanLehn, K. (2011). An analysis of students' gaming behaviors in an intelligent tutoring system: Predictors and impacts. *User Modeling and User-Adapted Interaction, 21*(1–2), 99–135.
Murphy, K. P. (2012). *Machine learning: A probabilistic perspective*. MIT press.
Natividad, G., Spector, J. M., & Evangelopoulos, N. (2018). *An analysis of two decades of educational technology publications: Who, what and where*. Springer.
Nielsen, R. D., Ward, W., & Martin, J. H. (2009). Recognizing entailment in intelligent tutoring systems. *Natural Language Engineering, 15*(4), 479–501.
Nie, D., Trullo, R., Lian, J., Petitjean, C., Ruan, S., Wang, Q., et al. (2017). Medical image synthesis with context-aware generative adversarial networks. In *International conference on medical image computing and computer-assisted intervention* (pp. 417–425). Springer.
Ninaus, M., Greipl, S., Kiili, K., Lindstedt, A., Huber, S., Klein, E., et al. (2019). Increased emotional engagement in game-based learning–A machine learning approach on facial emotion detection data. *Computers & Education, 142*, 103641.
Ocaña-Fernández, Y., Valenzuela-Fernández, L. A., & Garro-Aburto, L. L. (2019). Artificial intelligence and its implications in higher education. *Journal of Educational Psychology-Propositos y Representaciones, 7*(2), 553–568.

Ohba, N., Nakao, K., Isashiki, Y., & Ohba, A. (2007). The 100 most frequently cited articles in ophthalmology journals. *Archives of Ophthalmology, 125*(7), 952–960.

Olshausen, B. A., & Field, D. J. (1996). Emergence of simple-cell receptive field properties by learning a sparse code for natural images. *Nature, 381*(6583), 607.

Park, K., Mott, B. W., Min, W., Boyer, K. E., Wiebe, E. N., & Lester, J. C. (2019). Generating educational game levels with multistep deep convolutional generative adversarial networks. In *2019 IEEE conference on games (CoG)* (pp. 1–8). IEEE.

Peters, M. A. (2018). *Deep learning, education and the final stage of automation*. Taylor & Francis.

Pierre-Yves, O. (2003). The production and recognition of emotions in speech: Features and algorithms. *International Journal of Human-Computer Studies, 59*(1–2), 157–183.

Piramuthu, S. (2005). Knowledge-based web-enabled agents and intelligent tutoring systems. *IEEE Transactions on Education, 48*(4), 750–756.

Pliakos, K., Joo, S.-H., Park, J. Y., Cornillie, F., Vens, C., & Van den Noortgate, W. (2019). Integrating machine learning into item response theory for addressing the cold start problem in adaptive learning systems. *Computers & Education, 137*, 91–103.

Poole, D. I., Goebel, R. G., & Mackworth, A. K. (1998). *Computational intelligence and knowledge*. New York: Oxford University Press.

Rodrigues, F., & Oliveira, P. (2014). A system for formative assessment and monitoring of students' progress. *Computers & Education, 76*, 30–41.

Rojas, E. M., & Mukherjee, A. (2006). Multi-agent framework for general-purpose situational simulations in the construction management domain. *Journal of Computing in Civil Engineering, 20*(3), 165–176.

Roll, I., & Wylie, R. (2016). Evolution and revolution in artificial intelligence in education. *International Journal of Artificial Intelligence in Education, 26*(2), 582–599.

Romero, C., & Ventura, S. (2007). Educational data mining: A survey from 1995 to 2005. *Expert Systems with Applications, 33*(1), 135–146.

Romero, C., & Ventura, S. (2010). Educational data mining: A review of the state of the art. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews), 40*(6), 601–618.

Romero, C., & Ventura, S. (2013). Data mining in education. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 3*(1), 12–27.

Romero, C., Ventura, S., & García, E. (2008). Data mining in course management systems: Moodle case study and tutorial. *Computers & Education, 51*(1), 368–384.

Romero, C., Zafra, A., Luna, J. M., & Ventura, S. (2013). Association rule mining using genetic programming to provide feedback to instructors from multiple-choice quiz data. *Expert Systems, 30*(2), 162–172.

Russell, S. J., & Norvig, P. (2016). *Artificial intelligence: A modern approach*. Malaysia: Pearson Education Limited.

Samuel, A. (1959). Some studies in machine learning using the game of checkers. *IBM Journal of Research and Development, 3*(3), 210–229.

Schmidhuber, J. (2015). Deep learning in neural networks: An overview. *Neural Networks, 61*, 85–117.

Siemens, G., & Baker, R. S. D. (2012). Learning analytics and educational data mining: Towards communication and collaboration. In *Proceedings of the 2nd international conference on learning analytics and knowledge* (pp. 252–254). ACM.

Song, Y., Chen, X., Hao, T., Liu, Z., & Lan, Z. (2019). Exploring two decades of research on classroom dialogue by using bibliometric analysis. *Computers & Education, 137*, 12–31.

Suchman, L. A. (1987). *Plans and situated actions: The problem of human-machine communication*. Cambridge university press.

Sullivan, F. R., & Keith, P. K. (2019). Exploring the potential of natural language processing to support microgenetic analysis of collaborative learning discussions. *British Journal of Educational Technology, 50*(6), 3047–3063.

Sweileh, W. M., Al-Jabi, S. W., Zyoud, S.e. H., & Sawalha, A. F. (2018). Bibliometric analysis of literature in pharmacy education: 2000–2016. *International Journal of Pharmacy Practice, 26*(6), 541–549.

Tang, S., Peterson, J. C., & Pardos, Z. A. (2016). Deep neural networks and how they apply to sequential education data. In *Proceedings of the third (2016) ACM conference on learning@ scale* (pp. 321–324).

Sze, V., Chen, Y. H., Yang, T. J., & Emer, J. S. (2017). Efficient processing of deep neural networks: A tutorial and survey. *Proceedings of the IEEE, 105*(12), 2295–2329.

Tan, D.-P., Ji, S.-M., & Jin, M.-S. (2012). Intelligent computer-aided instruction modeling and a method to optimize study strategies for parallel robot instruction. *IEEE Transactions on Education, 56*(3), 268–273.

Tassopoulos, I. X., & Beligiannis, G. N. (2012). Solving effectively the school timetabling problem using particle swarm optimization. *Expert Systems with Applications, 39*(5), 6029–6040.

Tatar, D., Roschelle, J., Vahey, P., & Penuel, W. R. (2003). Handhelds go to school: Lessons learned. *Computer*, (9), 30–37.

Thomaz, A. L., & Breazeal, C. (2008). Teachable robots: Understanding human teaching behavior to build more effective robot learners. *Artificial Intelligence, 172*(6–7), 716–737.

Tran, B. X., Latkin, C. A., Vu, G. T., Nguyen, H. L. T., Nghiem, S., Tan, M.-X., et al. (2019). The current research landscape of the application of artificial intelligence in managing cerebrovascular and heart diseases: A bibliometric and content analysis. *International Journal of Environmental Research and Public Health, 16*(15), 2699.

Tsai, C. J., Tseng, S. S., & Lin, C. Y. (2001). May). A two-phase fuzzy mining and learning algorithm for adaptive learning environment. In *International Conference on Computational Science* (pp. 429–438). Berlin, Heidelberg: Springer.

Tsiriga, V., & Virvou, M. (2004). A framework for the initialization of student models in web-based intelligent tutoring systems. *User Modeling and User-Adapted Interaction, 14*(4), 289–316.

Verdú, E., Verdú, M. J., Regueras, L. M., de Castro, & García, R. (2012). A genetic fuzzy expert system for automatic question classification in a competitive learning environment. *Expert Systems with Applications, 39*(8), 7471–7478.

Wang, F. H. (2017). An exploration of online behaviour engagement and achievement in flipped classroom supported by learning management system. *Computers & Education, 114*, 79–91.

Wei, X., Lin, H., Yang, L., & Yu, Y. (2017). A convolution-LSTM-based deep neural network for cross-domain MOOC forum post classification. *Information, 8*(3), 92.

Whitehill, J., Serpell, Z., Lin, Y. C., Foster, A., & Movellan, J. R. (2014). The faces of engagement: Automatic recognition of student engagementfrom facial expressions. *IEEE Transactions on Affective Computing, 5*(1), 86–98.

Xu, D., & Wang, H. (2006). Intelligent agent supported personalization for virtual learning environments. *Decision Support Systems, 42*(2), 825–843.

Zawacki-Richter, O., & Latchem, C. (2018). Exploring four decades of research in Computers & Education. *Computers & Education, 122*, 136–152.

Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education–where are the educators? *International Journal of Educational Technology in Higher Education, 16*(1), 39.

Zhao, H., Li, G., & Feng, W. (2018). Research on application of artificial intelligence in medical education. In *2018 international conference on engineering simulation and intelligent control (ESAIC)* (pp. 340–342). IEEE.

Zhao, L., Zhou, Y., Lu, H., & Fujita, H. (2019). Parallel computing method of deep belief networks and its application to traffic flow prediction. *Knowledge-Based Systems, 163*, 972–987.

Zhong, Q., Li, B.-H., Zhu, Q.-Q., Zhang, Z.-M., Zou, Z.-H., & Jin, Y.-H. (2019). The top-100 highly cited original articles on immunotherapy for childhood leukemia. *Frontiers in Pharmacology, 10*, 1100.

Zouaq, A., & Nkambou, R. (2008). Building domain ontologies from text for educational purposes. *IEEE Transactions on learning technologies, 1*(1), 49–62.

Zovko, V., & Gudlin, M. (2019). Artificial intelligence as a disruptive technology in education. In *9th international conference the future of education*.